

FIGURE 1

ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA
CCCACGCGTCCGGGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCTCTTCCCG
CAGCGCTACCCGCCATGCGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTG
CTGCTGCCGCCCCGCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCT
GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACA
CGGCTTGGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTGCGCTGCTGGAGATC
CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA
GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC
CAGGGCGGATCCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA
GGGCGACGGGTCTGCCGGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGG
ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC
TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT
GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTG
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACAGTGCAGAGAGTGTGACTCCAGCTGTGTG
GGCTGCACAGGGGAAGGCCCAGGAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA
CGGACAGTGTGCAGATGTGGACGAGTGTCTACTAGCAGAAAAAACCTGTGTGAGGAAAAACG
AAAAGTGTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCTGACGGCTTCGAAGAAACG
GAAGATGCCTGTGTGCCGCCGCGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT
GCCCTCCCGCGAAGACCTGTAATGTGCCGGACTTACCCTTTAAATTATTCAGAAGGATGTCC
CGTGGAATAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGC
CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA
AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG
AGGAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGC
TTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA
TCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTC
ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTTCGGCGCAGCACCATGGCCTGAAAT
AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG
TGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAGCATGC
ATCTCAATTAGTCAGCAACCCAGTTTTT

U950011.1 "U950011.1"

FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop

><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEKTLISKYESSSEIRL
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKVCCSPGTYGPDCLACQGGSQRPCSG
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESKTCSGLTNRDCGECEVGWVLDE
GACVDVDECAAEPPPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKT
CVRKNENCYNTPGSYVCVCPDGFEEETEDACVPPAEAEATEGESPTQLPSREDL

Signal peptide:

amino acids 1-24

N-glycosylation sites.

amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.

amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 26-30

Casein kinase II phosphorylation sites.

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343
and 349-353

Tyrosine kinase phosphorylation site.

amino acids 303-310

N-myristoylation sites.

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and
313-319

Aspartic acid and asparagine hydroxylation site.

amino acids 308-320

EGF-like domain cysteine pattern signature.

amino acids 166-178

Leucine zipper pattern.

amino acids 94-116

FIGURE 3

U550401.1 "U550401.1"

CAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTC
GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCGCCAGCCGTCTAAACGGGAACA
GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG
GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGCATGGCCCCGGAGGAGCGCCTTC
CCTGCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGC
CGGGCCGCCGAGGAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA
TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGAT
TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCTGTCAATATCCATTCCATGAATTTTAC
CTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGAATTCCTGTCTTGGCGCTCCCTGGATA
AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA
TCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTGAAAGT
GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT
TCTTTAAACATGTCAACAAGCTGAGTGCCCGAGGCGGGTGCCGAAATGGAGGCTTTTGTAAT
GAAAGACGCATCTGCGAGTGTCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTG
TACCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTG
GATTCTATGGAGTGAACGTGACAAAGCAAATGCTCAACCACCTGCTTTAATGGAGGGACC
TGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG
CAAATGCCCAACCCCTGTGCAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT
CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT
GGAACCTGCCATGAACCCAACAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAA
TAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCGCCAGCTCAGGCAGC
ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG
TGAACCTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTAAACCTTTCA
TGTGTTGAATGTTCAAATAATGTTTATTACACTTAAGAATACTGGCCTGAATTTTATTAGCT
TCATTATAAATCACTGAGCTGATATTTACTCTTCTTTTAAAGTTTTCTAAGTACGTCTGTAG
CATGATGGTATAGATTTTCTTGTTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA
TCAGGTTAAAATTTTTCAGTGTGTAGTTGGCAGATATTTTCAAAATTACAATGCATTTATGGT
GTCTGGGGGCAGGGGAACATCAGAAAGGTTAAATTTGGGCAAAAATGCGTAAGTCACAAGAAT
TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGAT
GTTTGTACATTTTAAAAAATTGCTCTTAATTTTAACTCTCAATACAATATATTTTGACC
TTACCATTATTCCAGAGATTGAGTATTAATAAAAAAAAAAATTACACTGTGGTAGTGGCATT
AAACAATATAATATATTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCAT
TGGCTTGAAGCAATATAATATATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTAT
ACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTTTTGGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGACTCTAGAGTCGACCTGCAGAAGCTTGGC
CGCCATGGCCCACTTGTTTATTGCAGCTTATAATG

FIGURE 4

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094

><subunit 1 of 1, 379 aa, 0 stop

><MW: 41528, pI: 7.97, NX(S/T): 2

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESELYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPPGGC
RNGGFCNERRICECPDGFHGPHEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNNGGTCTFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAERRDP
PESNYIW

Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 88-92, 245-249

Casein kinase II phosphorylation site.

amino acids 319-323

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

FIGURE 6

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223

><subunit 1 of 1, 164 aa, 1 stop

><MW: 18359, pI: 7.45, NX(S/T): 1

MWRCPLGLLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQLCCRGRADDC
ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPFPPIQGCMHGGRIYPVLGTYWDNCNR
CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,
411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,
269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 398-409

FIGURE 7

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTTCGTTGTCTCAAATCCA
ATTCTCTTGGGACACATNACGCCTGTCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCTTGCCTCGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
CCCCCTGTATGATGCACAGCCGAGCCATGGGTCTGGGGCAAGCGCCAGGCCACTGCCCCTGCGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG
TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG
AGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

0550401-07101
TCTG

FIGURE 8

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG
GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG
CAGCTTGAAGTGCCTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT
GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT
GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG
GCCCAGTGGGAGCCTGTCTTGGTTCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT
CTGACCCCCTGTCCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTC
CATGGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT
CCCTGCCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA
GGGGACAGGCACTCAGGAGGGGCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA
GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA
GGGGCCAGGCCTCACATTCTGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT
AATAAACACCTGTTGGATAAGCCAAAAAA

05504011-071101

DATE	TIME	LOCATION	WIND	WAVE	SEA	TEMP	WIND	WAVE	SEA	TEMP
11/11/54	0800	1000	10	10	10	10	10	10	10	10
11/11/54	0900	1000	10	10	10	10	10	10	10	10
11/11/54	1000	1000	10	10	10	10	10	10	10	10
11/11/54	1100	1000	10	10	10	10	10	10	10	10
11/11/54	1200	1000	10	10	10	10	10	10	10	10
11/11/54	1300	1000	10	10	10	10	10	10	10	10
11/11/54	1400	1000	10	10	10	10	10	10	10	10
11/11/54	1500	1000	10	10	10	10	10	10	10	10
11/11/54	1600	1000	10	10	10	10	10	10	10	10
11/11/54	1700	1000	10	10	10	10	10	10	10	10
11/11/54	1800	1000	10	10	10	10	10	10	10	10
11/11/54	1900	1000	10	10	10	10	10	10	10	10
11/11/54	2000	1000	10	10	10	10	10	10	10	10
11/11/54	2100	1000	10	10	10	10	10	10	10	10
11/11/54	2200	1000	10	10	10	10	10	10	10	10
11/11/54	2300	1000	10	10	10	10	10	10	10	10
11/11/54	2400	1000	10	10	10	10	10	10	10	10
11/11/54	0100	1000	10	10	10	10	10	10	10	10
11/11/54	0200	1000	10	10	10	10	10	10	10	10
11/11/54	0300	1000	10	10	10	10	10	10	10	10
11/11/54	0400	1000	10	10	10	10	10	10	10	10
11/11/54	0500	1000	10	10	10	10	10	10	10	10
11/11/54	0600	1000	10	10	10	10	10	10	10	10
11/11/54	0700	1000	10	10	10	10	10	10	10	10
11/11/54	0800	1000	10	10	10	10	10	10	10	10
11/11/54	0900	1000	10	10	10	10	10	10	10	10
11/11/54	1000	1000	10	10	10	10	10	10	10	10
11/11/54	1100	1000	10	10	10	10	10	10	10	10
11/11/54	1200	1000	10	10	10	10	10	10	10	10
11/11/54	1300	1000	10	10	10	10	10	10	10	10
11/11/54	1400	1000	10	10	10	10	10	10	10	10
11/11/54	1500	1000	10	10	10	10	10	10	10	10
11/11/54	1600	1000	10	10	10	10	10	10	10	10
11/11/54	1700	1000	10	10	10	10	10	10	10	10
11/11/54	1800	1000	10	10	10	10	10	10	10	10
11/11/54	1900	1000	10	10	10	10	10	10	10	10
11/11/54	2000	1000	10	10	10	10	10	10	10	10
11/11/54	2100	1000	10	10	10	10	10	10	10	10

Signal peptide:

N-glycosylation site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Amidation site.

amino acids 70-74

[illegible]

CCACACGCGTCCGAACCTCTCCAGCGATGGGGAGCCGCCCGCCTGCTGCCCCAACCTCACTCTGT
GCTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGCGCCATGACC
GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA
CGTGCAGGTCACCGGGCGTCCGATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC
TCATAGTGGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGCTGAGAGTGAGAAG
TACATCTGTATGAACAAGAGGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTG
CGTGTTACGGAGATCGTGCTGGAGAACAACCTATACGGCCTTCCAGAACGCCCCGGCACGAGG
GCTGGTTCATGGCCTTCACGCGGCAGGGGCGGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAAC
CAGCGCGAGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCCACGC
CGAGAAGCAGAAGCAGTTCGAGTTTGTGGGCTCCGCCCCCACCCGCCGGACCAAGCGCACAC
GGCGGCCCCAGCCCCCTCACGTAGTCTCTGGGAGGCAGGGGGCAGCAGCCCCCTGGGCCGCCTCCC
CACCCCTTTCCCTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCC
GAGGGAGGACCCTGAGGGCCGCGAAGCATCCGAGCCCCCAGCTGGGAAGGGGCAGGCCGGTG
CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT
GAGTGTCACCCTGATCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCC
CGCTGAAAGGTCAGCGACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCCTCAAAA
TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCCAGCCCCCAAACCTCCTCCTGGCTAGACTGTA
GGAAGGGACTTTTGTTTGTTTGTTCAGGAAAAAAGAAAGGGAGAGAGAGGAAAAATAG
AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCCACCCCCAACTCCCAGCCC
CGGAATAAAACCATTTCCTGC

FIGURE 11

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRISATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLENNYTAFQONARHEGWFMATRQGRPRQASRSRQONQREAHFIKRLYQGQLPFPNHAEKQKQFEFVGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 12

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG
GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT
CATTTTGATTTTGCTGTTTATTTTTTTTTTTCTTTTCTTTTCCCACCACATTGTATTTTAT
TTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT
GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAACTCCTGGCCTGCC
CTAGTGTGTGCCGCTGCGACAGGAACCTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC
TGGATTTCTCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC
AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC
AATATTCAGACCATTTACAGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCT
GGATGACAACCTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCC
TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCTGTGGAC
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCAATATCCGACATGGCCTTCCAGAA
TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG
AGGGCACCTTCAGCCATCTCACCAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC
CACCTCCTCCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT
AAACCACATTCTTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA
ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC
ACTGCTCGGAATAACCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAA
ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGG
GGATGGCCGTCAGGGAATTAATATGAATCTTTTGTCTGTCCCACCACGACCCCCGGCCTG
CCTCTCTTACCCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT
TCCAAACCCTAGCAGAAGCTACACGCCTCCAACCTCCTACCACATCGAACTTCCCACGATTC
CTGACTGGGATGGCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATC
CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTACCCTGATGGCATA
CAAACCTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTCAGGAGCGCATAG
TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT
TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGC
CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA
CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT
GTGCTGGTGGTCTTGCTCAGCGTCTTTTGTGCTGGCATATGCACAAAAGGGGCGCTACACCTC
CCAGAAGTGGAAATACAACCGGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA
AGGACAACCTCCATCCTGGAGATGACAGAAACAGTTTTTCAGATCGTCTCCTTAAATAACGAT
CAACTCCTTAAAGGAGATTTAGACTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTA
CACAGACTGCCATATCCCCAACACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC
ACTGCCATACGTGACAGCCAGAGGCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT
TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTG
CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTTAAATCTT

U.S. Patent & Trademark Office

[illegible]

Signal peptide:

Transmembrane domain:

N-glycosylation site.

Casein kinase II phosphorylation site.

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

FIGURE 14

U55441.1 "U55441.1"

ACTTGGAGCAAGCGGCGGCGGCGGAGACAGAGGCAGAGGCAGAAAGCTGGGGCTCCGTCTCGCCTCCACGAGCG
ATCCCCGAGGAGAGCCGCGGCCCTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCCCTGCC
TCGCTTCCCAGGCGCGGCGGCTGCAGCCTTGCCCCCTCTTGCTCGCCTTGAAAATGGAAAAGATGCTCGCAGGCT
GCTTTCTGCTGATCCTCGGACAGATCGTCCTCCTCCCTGCCGAGGCCAGGGAGCGGTACGCTGGGAGGTCCATCT
CTAGGGGCAGACACGCTCGGACCCACCCGAGACGGCCCTTCTGGAGAGTTCTGTGAGAACAGCGGGCAGACC
TGGTTTTTCATCATTTGACAGCTCTCGCAGTGTCAACACCCATGACTATGCAAAGGTCAAGGAGTTTCATCGTGGACA
TCTTGCAATTCTTGGACATTGGTCTGTATGTCAACCGAGTGGGCCTGCTCCAATATGGCAGCACTGTCAAGAATG
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGTGTCAAGAGGATGCGGCATCTGTCCACGG
GCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCGCGGCCCTGA
GGGAGAATGTGCCACGGGTCATAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA
AGGCACGGGACACGGGCATCCTAATCTTTGCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATTG
GGAGTGAGCCCCATGAGGACCATGTCTTCTTGTGGCCAATTTAGCCAGATTGAGACGCTGACCTCCGTGTTCC
AGAAGAAGTTGTGCACGGCCACATGTGCAGCACCTTGGAGCATAACTGTGCCACTTCTGCATCAACATCCCTG
GCTCATACGTCTGCAGGTGCAAAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGT
GTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTG
GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCTCAGAAAACACGGATGTGAAC
ATGAGTGTGTAAATGCTGATGGCTCCTACCTTTGCCAGTGCCATGAAGGATTTGCTCTTAACCCAGATGAAAAAA
CGTGACAAAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGCGTCAACATGGAGGAGAGCT
ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCAC
AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCTTCGTCTGCCAGTGCTCAGAAGGCTTCC
TCATCAACGAGGACCTCAAGACCTGCTCCCGGTGGATTACTGCCTGTGAGTGACCATGGTTGTGAATACTCCT
GTGTCAACATGGACAGATCCTTTGCCCTGTCACTGTCTGAGGGACACGTGCTCCGACGATGGGAAGACGTGTG
CAAAATTGGACTCTTGTGCTCTGGGGGACCACGGTTGTGAACATTCTGTGTGAAGCAGTGAAGATTCTGTTGTG
GCCAGTGCTTTGAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAAGATGTCTGCCAAGCTATAG
ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACAGTGGAGTGCTTGGAGGGATTCCGGCTCG
CTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAATCAACCCACCATGGCTGCCAACACATTTGTGTTA
ATAATGGGAATTCCTACATCTGCAATGCTCAGAGGGATTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT
GCACCTGAAGGCCCAATTGACCTGGTCTTTGTGATCGATGGATCCAGAGTCTTGAGAGAAGAGAATTTTGAGGTG
TGAAGCAGTTTGTCACTGGAATTATAGATTCTTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGT
ATTCCACACAGGTCCACACAGAGTTCACTCTGAGAACTTCAACTCAGCCAAAGACATGAAAAAAGCCGTGGCCC
ACATGAAATACATGGGAAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTACCCAAG
GAGAAGGGGCCAGGCCCCCTTCCACAAGGGTGCCAGAGCAGCCATTGTGTTACCGACGACGGGCTCAGGATG
ACGTCTCCGAGTGGGCCAGTAAAGCCAAGGCCAATGGTATCACTATGTATGCTGTTGGGGTAGGAAAAGCCATTG
AGGAGGAATAACAAGAGATTGCCTCTGAGCCCAACAAACAAGCATCTTCTATAGCCGAAGACTTCAGCACAAATG
ATGAGATAAGTGAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG
CAGGGGAACCTGCCAAAAACGGTCCAACAGCCAAACAGAATCTGAGCCAGTCACCATAAATATCCAAGACCTACTTT
CCTGTTCTAATTTTGCAGTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTT
CCCATTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAAACACGATCAATGCAAATGTGAAAACTTATAATGT
TCCAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC
TGGAAAATCGCCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCATTGTATCACGATTACAAT
GAACGCAGTGACAGAGCCCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAAAACAATCAGTACTGA
GAAACCTGGTTTGGCACAGAACAAGACAAGAAGTATACACTAAGTTGTATAAATTTATCTAGGAAAAAAATCCT
TCAGAATTCTAAGATGAATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAAATATACTGTGGACAC
AACTTGCTTCTGCCTCATCTGCCTTAGTGTGCAATCTCATTTGACTATACGATAAAGTTTGCACAGTCTTACTT
CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTGTACTGGACTTTACCTTGATATATGTATATGGATGTATG
CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAATAATTCACCACTTCAG

FIGURE 15

MEKMLAGCFLILGQIVLLPAEARERSRGRSISRGRHARTHPTALLESSCENKRADLVFII
DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV
KMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD
TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN
CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA
LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGC
EHECVNMEESYYCRCHRGYTLDPNGKTC SRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLI
NEDLKTC SRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTC AKLDSCALGDHGCE
HSCVSSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLA
EDGKR¹RRKDVC²KSTHHGCEHICVNNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVID
GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVA
HMKYMGKGSMTGLALKHMFERSFTQEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN
GITMYAVGVGKAIEEELQEIASEPTNKHLYAEDFSTMDIEISEKLKKGICEALEDSDGRQDS
PAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL
EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247,
401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784,
781-785, 819-823, 866-870

N-myristoylation site.

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500,
639-645, 690-696, 752-758, 792-798

Amidation site.

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464,
540-546, 581-587

FIGURE 15

FIGURE 16

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTGCGCGAGCCTCG
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGGCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGCGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCCGACCCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCCT
TTCCTCTGTACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT
TGACCCCCACGCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT
GGGGCCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTTCTTCCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCAGAGTCAAGCTGGCTCTCTGGTATGATGAC
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCTATAAGGGTCACCTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCAGGC
CCCCTTGATCTGTACCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAAAACTAACATGAAATATGTGTTGTTTTCATTTGCAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

U01704560

FIGURE 17

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS
SQPWEVPPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY
SCSVNVQDKQGKSRGHSIKTLELNLVLPAPPSCRLQGVPHVGANVTLSQCSPRSKPAVQYQ
WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA
VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPWKSSDTISKNGTL
SSVTSARALRPPHGP RP GALTP T PSLSSQALPSPRLPTTDGAHPQPI SPIPGGVSSSGLSR
MGAVPVMVPAQSQAGSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

FIGURE 19

MKRLPLL VVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC
GNLTQSCGENANCTNTEGSYYCMCVPGFRSSNQDRFITNDGTVCIENVNANCHLDNVCIAA
NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSLLGYKNNTISAKDTL
SNSTLTETVKT VNNFVQRDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT
NSTDIALKVFFFDSYNMKHIHPHMNMDGDYINIFPKRKAAYDSNGNVAVAFLYYK SIGPLLS
SSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTD RYRSLCAF
WNYSPTMNGSWSSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG
IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFC SIIAGL
LHYFFLAFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLSPAVVVGFS AALGYRYYGT
TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA
LALLFLLGTTWIFGVLHVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNV
PCCFGCLR

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636,
648-664

N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181,
188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154,
155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329,
346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394,
434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

FIGURE 19

1000 999 998 997 996 995 994 993 992 991 990 989 988 987 986 985 984 983 982 981 980 979 978 977 976 975 974 973 972 971 970 969 968 967 966 965 964 963 962 961 960 959 958 957 956 955 954 953 952 951 950 949 948 947 946 945 944 943 942 941 940 939 938 937 936 935 934 933 932 931 930 929 928 927 926 925 924 923 922 921 920 919 918 917 916 915 914 913 912 911 910 909 908 907 906 905 904 903 902 901 900 899 898 897 896 895 894 893 892 891 890 889 888 887 886 885 884 883 882 881 880 879 878 877 876 875 874 873 872 871 870 869 868 867 866 865 864 863 862 861 860 859 858 857 856 855 854 853 852 851 850 849 848 847 846 845 844 843 842 841 840 839 838 837 836 835 834 833 832 831 830 829 828 827 826 825 824 823 822 821 820 819 818 817 816 815 814 813 812 811 810 809 808 807 806 805 804 803 802 801 800 799 798 797 796 795 794 793 792 791 790 789 788 787 786 785 784 783 782 781 780 779 778 777 776 775 774 773 772 771 770 769 768 767 766 765 764 763 762 761 760 759 758 757 756 755 754 753 752 751 750 749 748 747 746 745 744 743 742 741 740 739 738 737 736 735 734 733 732 731 730 729 728 727 726 725 724 723 722 721 720 719 718 717 716 715 714 713 712 711 710 709 708 707 706 705 704 703 702 701 700 699 698 697 696 695 694 693 692 691 690 689 688 687 686 685 684 683 682 681 680 679 678 677 676 675 674 673 672 671 670 669 668 667 666 665 664 663 662 661 660 659 658 657 656 655 654 653 652 651 650 649 648 647 646 645 644 643 642 641 640 639 638 637 636 635 634 633 632 631 630 629 628 627 626 625 624 623 622 621 620 619 618 617 616 615 614 613 612 611 610 609 608 607 606 605 604 603 602 601 600 599 598 597 596 595 594 593 592 591 590 589 588 587 586 585 584 583 582 581 580 579 578 577 576 575 574 573 572 571 570 569 568 567 566 565 564 563 562 561 560 559 558 557 556 555 554 553 552 551 550 549 548 547 546 545 544 543 542 541 540 539 538 537 536 535 534 533 532 531 530 529 528 527 526 525 524 523 522 521 520 519 518 517 516 515 514 513 512 511 510 509 508 507 506 505 504 503 502 501 500 499 498 497 496 495 494 493 492 491 490 489 488 487 486 485 484 483 482 481 480 479 478 477 476 475 474 473 472 471 470 469 468 467 466 465 464 463 462 461 460 459 458 457 456 455 454 453 452 451 450 449 448 447 446 445 444 443 442 441 440 439 438 437 436 435 434 433 432 431 430 429 428 427 426 425 424 423 422 421 420 419 418 417 416 415 414 413 412 411 410 409 408 407 406 405 404 403 402 401 400 399 398 397 396 395 394 393 392 391 390 389 388 387 386 385 384 383 382 381 380 379 378 377 376 375 374 373 372 371 370 369 368 367 366 365 364 363 362 361 360 359 358 357 356 355 354 353 352 351 350 349 348 347 346 345 344 343 342 341 340 339 338 337 336 335 334 333 332 331 330 329 328 327 326 325 324 323 322 321 320 319 318 317 316 315 314 313 312 311 310 309 308 307 306 305 304 303 302 301 300 299 298 297 296 295 294 293 292 291 290 289 288 287 286 285 284 283 282 281 280 279 278 277 276 275 274 273 272 271 270 269 268 267 266 265 264 263 262 261 260 259 258 257 256 255 254 253 252 251 250 249 248 247 246 245 244 243 242 241 240 239 238 237 236 235 234 233 232 231 230 229 228 227 226 225 224 223 222 221 220 219 218 217 216 215 214 213 212 211 210 209 208 207 206 205 204 203 202 201 200 199 198 197 196 195 194 193 192 191 190 189 188 187 186 185 184 183 182 181 180 179 178 177 176 175 174 173 172 171 170 169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154 153 152 151 150 149 148 147 146 145 144 143 142 141 140 139 138 137 136 135 134 133 132 131 130 129 128 127 126 125 124 123 122 121 120 119 118 117 116 115 114 113 112 111 110 109 108 107 106 105 104 103 102 101 100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

TGGAAACATATCCTCCCTCATATGAATATGGATGGAGACTACATAAAATATATTTCCAAAGNG
AAAAGCCGGCATATGGATTCAAATGGCAATGTTGCAGTTGCATTTTATATTATAAGAGTAT
TGGTCCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAATATTATGATAATTCT
GAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTAGTCTCAATGAGCTCAAACCCACCCAC
ATTATATGAACTTGAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATAGGTATA
GGAGTCTATGTGGCATTTTTGAATACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAG
GGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACA
TTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTCTTACAA
GGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTTTACCTTC
TGGTTCTTCAGTGAAATTCAAAGCACCAGGA

FIGURE 21

GCTCCCAGCCAAGAACCTCGGGGCGGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG
CTAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTACCCCGGCTGGGA
CAAGAAGCCGCGCCTGCCTGCCCGGGCCCGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGG
GGTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG
CCGCCGCGCCGTACAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC
GGCACTCACAGCCCCGCAGCGCATCCCGGTGCGCGCCAGCCTCCCGCACCCCATCGCCGG
AGCTGCGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGGTCCACGTATGG
ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC
CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCC
ACGGGCTCTCCAGCTGCTTCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGGC
CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG
CGTGACAGCGTGCGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT
ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA
TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA
CAGAGGCTTTCTTCCACTCTCTCATTTCTGCCCCATGCTGCCCATGGTCCCAGAGGAGCCTG
AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCCCTGGAGACCGACAGCATG
GACCCATTTGGGCTTGTACCGGACTGGAGGCCGTGAGGAGTCCCAGCTTTGAGAAGTAACT
GAGACCATGCCCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG
TGCTTCTACAAGAACAGTCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA
GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT
CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGCCTGGGCCCCCATTTCTGCTCCCTCGA
GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC
TCACTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC
CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTTCAGGAACAGGTGATCCACTCTGTA
AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
ACCATTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCACCAG
GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACCTTGAGAATTCCCC
CTGAGGCCAGTTCTGTCTATGGATGCTGTCTGAGAATAAATTGCTGTCCCGGTGTCACCTGC
TTCCATCTCCCAGCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCCATGGATTGGGGCCT
CCCAGGCCCCCACCCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT
TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC
CCTTTCCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACA
TCTCTTATTTTCTTACATTATTTATGCCCCCAAATTATTTATGTATGTAAGTGAGGTTTG
TTTTGTATATTAAATGGAGTTTGTGTGT

U.S. Patent & Trademark Office

1. The first part of the paper is devoted to the study of the properties of the function $f(x)$ defined by the equation $f(x) = \sum_{n=0}^{\infty} a_n x^n$, where a_n are the coefficients of the power series.

Signal peptide:

Casein kinase II phosphorylation site.

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

FIGURE 23

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA
GAGCAGCCGGCTGCCGCCCCGGGAAGATGGCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTG
CTGCTGCGCTACCTGGTGGTTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA
AGAAGACTGTTTTCTCCAGATTAGAGTGGAAGAACTGGGTTCGGAGTGTCTCCTTTGTCTAC
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG
AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT
CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAACGTGGGTAGAGCTACGATGTCAAGA
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATAACAATGAATACAAAACTGGAAC
CTGCAATTTAATACTGTTTCCAACTGGACACTGGAGAATATTCCCTGTGAAGCCCGCAATTC
TGTTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT
GCTCAGAGGAAAGGCTACTTTTCAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA
AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG
CCGCGGCGGGCGGATCACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAATATGGTGAAACCC
CATCTCTACTAAAATACAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGC
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC
CACTGCAGTCCAGCCTGGGTAACAGAGCAAGATTCCATCTCAAAAATAAAATAAATAAATA
AATAAATACTGGTTTTTACCTGTAGAATTCTTACAATAAATATAGCTTGATATTC

U
T
M
E
T
H
O
D
S

[illegible]

amino acids 1-20

amino acids 130-144, 238-258

amino acids 98-102, 187-191, 236-240, 277-281

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

amino acids 226-230

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	

AAAAACCACGAAGGAACTACTCCAAAATGAAC

[illegible]

Signal sequence:

amino acids 277-280

УДК 62-50

GCCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG
CTGCAGCCTTTTGAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC
CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT
GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTGTCTTA
TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG
GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCTCAATGAAATTTTTAAGG
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT
GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG
TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCAC
AACGTGATCTGTAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC
CAACGACGCTGACCTTTGTAACCTCCCTAAAAAACTACCGATTATGCCATGCTGGTCACCA
TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAATCAGGAG
GATGCCCCGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA
ACCTGATGATATTAGCACTGTGGTATTAGTGTCCAACTGACTGTCATTGAGAAAGAAAGAAA
GTAGTTTTCGATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA
ACTTTGTATTTTCAGTTTTTTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACA
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCGATGGTATATTTCTGAGT
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT
AATTTAAAAGCAAATAAAAGCTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAAAAACA

[illegible]

Signal sequence:

Transmembrane domain:

N-glycosylation site.

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

N-myristoylation site.

amino acids 37-43, 45-51, 110-116

[illegible]

ACCGAGCCGAGCGGACCGGAAGGCGCGCCCGAGATGACAGGTGAGCAAGAGGATGCTGGCGGG
 GGCGTGAGGAGCATGCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCTCTCTGTTGCT
 GGGCTCAGTGCTGTCAGGCTCGGCCACGGGCTGCCGCCCGCTGCGAGTGCTCCGCCCAGG
 ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGGCATCCCCACCGAG
 ACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAACGCTCAACCAGGACGAGTTTCGCCAG
 CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCG
 CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGAGCAACCGCCTGAAGCTCATC
 CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT
 CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG
 ACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG
 ACGCTGGAGAAATGCAACCTGACCTCCATCCCACCGAGGCGCTGTCCACCTGCACGGCCT
 CATCGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC
 TGTACCGACTCAAGGTCTTGAGATCTCCACTGGCCCTACTTGACACCATGACACCCAAC
 TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC
 CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCCTCAACCTCTCTACAACCCCATCA
 GCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGG
 GGGCAGCTGGCCGTGGTGGAGCCCTATGCCCTTCCGCGGCCTCAACTACCTGCGCGTGTCTCA
 TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG
 AGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTCGGCTCCTGTGGGTGTTCCGG
 CGCCGCTGGCGGTCAACTTCAACCGGCAGCAGCCACGTGCGCCACGCCCGAGTTTGTCCA
 GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCCGCCGCG
 CCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTGTGGACGAGGGCCACACGGTGCAGTTT
 GTGTGCCGGGCCGATGGCGACCCGCCGCCGCCATCCTCTGGCTCTCACCCCGAAAGCACCT
 GGTCTCAGCCAAGAGCAATGGGCGGTTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT
 ACGCCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACGCGGGCGGCAACGAC
 TCCATGCCCGCCACCTGCATGTGCGCAGCTACTCGCCGACTGGCCCCATCAGCCCAACAA
 GACCTTCGCTTTCATCTCCAACAGCCGGGCGAGGGAGAGGCCAACAGCACCCGCGCCACTG
 TGCCTTTCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC
 CTGGGCGTCTCTCTTCTGCTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC
 AAAGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGAAGCAGGCATCAGCTCCGCCG
 ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGGGCCGGGGCGGGGGGCAGGGACCCCCG
 GCGCGCCGGGCAGGGGAAGGGGCCTGGTGCACACTGCTCACTCTCCAGTCTTCCACCTC
 CTCCCTACCCTTCTACACACGTTCTCTTTCTCCCTCCGCTCCGTCCCTGCTGCCCCCG
 CCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCAGACCTGGGGACCCCA
 CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCA
 ATAATTCAATAAAAAAGTTACGAACTTTCTCTGTAACCTGGGTTTCAATAATTATGGATTTT
 TATGAAAACCTTGAATAATAAAAAAGAGAAAAAACTAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MQVSKRMLAGGVRSMPSPLLACWQPIILLVLGSLVSGSATGCPPRCECSAQDRAVLCHRKCF
VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFFNNLFNLRTL
GLRSNRLKLIPLGVFTGLSNLTQDISENKIVILLDYMFDLYNLKSLEVGDNDLVYISHRA
FSGLNSLEQLTLEKCNLTSLIPTALSHLHGLIVLRLRHLNINAI RDYSFKRLYLKVL EISH
WPYLDTMTPNCLYGLNLTSLITHCNLTAVPYLAVRHLVYL RFLNLSYNPISTIEG SMLHEL
LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLT TLEESVFH SVGNLETLILDSNPLA
CDCRLLWVFRRRWRLNFNRRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV
FVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGR LTVFPDGTLEVR YAQVQDNGTYL
CIAANAGGNDSMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI
IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPRKSDAGISSADAPRKFNMKMI

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 556-578

N-glycosylation site.

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345,
492-496, 505-509, 526-530, 542-546

Casein kinase II phosphorylation site.

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353,
607-611

Tyrosine kinase phosphorylation site.

amino acids 590-598

N-myristoylation site.

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143,
262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 32-43

Figure 1 displays 12 histograms, labeled x_0 through x_{11} , showing the distribution of the number of non-zero elements in the vector x_k . The x-axis represents the number of non-zero elements (0 to 10), and the y-axis represents the count (0 to 10). The distributions are roughly bell-shaped and centered around 5, with the peak count increasing from 10 at x_0 to 12 at x_{11} .

Figure 1 displays 12 histograms, labeled x_0 through x_{11} , showing the distribution of the number of non-zero elements in the vector x_k . The x-axis represents the number of non-zero elements (0 to 10), and the y-axis represents the count (0 to 10). The distributions are roughly bell-shaped and centered around 5, with the peak count increasing from 10 at x_0 to 12 at x_{11} .

FIGURE 32

MGAPAASLLLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSSLQW
SNPAQQTLYFGEKRALRDNRIQLVTSTPHELSSISINVALADEGEYTC SIFTMPVRTAKSLV
TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGK
TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQ
KLLHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKA
YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIIMLI FLGHYLIRHKGTYLTHEAKGSDD
APDADTAIINAEGGQSGGDDKKEYFI

Signal sequence:

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,
306-310, 334-340, 360-364, 385-389, 386-390

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

Downloaded from www.sciencedirect.com

FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCTTTTCTTCTCCTTTCCTGG
CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG
TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTT
GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA
TCTGTTCCCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAGGGCTTCACA
AGTCTGCAGCGTTTCACTGCCCGACTTCCCAGTTTTACCATTTATTTCTGCATGGCAATTC
CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG
AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGG
CTGCACATCAACAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA
TCTGGAATATCTCCAGGCTGATTTTAAATTTATTACGAGATATAGACCCGGGGGCTTCCAGG
ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCTACCTGCCAAC
GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC
CTATGAGGAGGTCTTGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCTT
GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC
CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC
CACCGAACAGGACTTGTGTCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG
CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG
GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT
CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCTTAGCTAACA
GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAAC
TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA
GCTTTTCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA
ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG
AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA
GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA
TCCTCCCGGGCACTTTCAATGCCATGCCCAAACCTGAGGATCCTCATTTCTCAACAACAACCTG
CTGAGGTCCCTGCCTGTGGACGTGTTTCGCTGGGGTCTCGCTCTCTAAACTCAGCCTGCACAA
CAATTACTTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA
TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA
GAACGCTTGGGTTCGGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACCTTCTT
TAGAAAGGATTTTCATGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT
CGCCCACGTAACTTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC
AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT
GTTTGTACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTATTATCCTGAGGAACCGAAAGC
GGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTCCCTACAGACAGTCTGT
GACTCTTCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCACAGAGTGTATGACTG
TGGCTCTCACTCGCTCTCAGACTAAGACCCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG
ATACATCCTTCCCCACCGCAGGCACCCCGGGGGCTGGAGGGGCGTGTACCCAAATCCCCGCG
CCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT
GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA
GCCAGCTCGCTCTTTGCTGAGAGCCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAG
AACTGACAGTGCCCTCGCCCTCGCCCCCGGGGCTGTGGGGTTGGATGCCGCGGTTCTATAC
ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTTCCCTGTGGATTAG
CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT
TGTAATAAGTAACTTTGACTTCTGAC

[illegible]

Signal sequence:

amino acids 337-348

FIGURE 35

U
S
P
A
T
E
N
T

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCTTGACCCCAGAATAACTCAGGGC
TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCTGTGCGGGCCTAAGGGAAACTGTTGGC
CGCTGGGCCCCGCGGGGGGATTCTTGGCAGTTGGGGGGTCCGTGCGGAGCGAGGGCGGAGGGG
AAGGGAGGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC
AGCTCTGCGTCTCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGCGGGGCCTCAG
AGAATGAGGCCGGCGTTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCCGGGCGGGCGG
CGGCGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTGC
ACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCGCTC
AGCACCGTGCGTGCGGGCGCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGG
GCCCCGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTGCGACTGGAGCGCAGGCGTTCCCACT
GCACCCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCTCCGACCCCGGCGGTCTC
GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGC
GGTACTCCAGGCCACCGGTGGGGTTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC
GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCCTGCGCCGCGCCCCGGG
GCCGCCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGCGCCGCTCTGGACTTCAG
TCCACCTGGGACCGAGGTGAGTGCCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA
TCGCGGACGAAATCGGCGCTCGCTGGGACAACTCTCGGGCGATGTGTTGTGTCCTGCCCC
GGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG
CTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA
GTGGGGAAGGACAGCCGACCTTGGGGGGACCGGGGTGCCACCAGGCGCCCGCGGCCACT
GCAACCAGCCCCGTGCCGAGAGAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC
ACCACTTGTCCCTGAACAAGACAATTTCAGTAACATCTATTCTGAGATTCTCGATGGGGAT
CACAGAGCACGATGTCTACCCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC
CCATCAGGGAGCGTGATTTCCAAGTTTAAATTCTACGACTTCTCTGCCACTCCTCAGGCTTT
CGACTCCTCCTCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTGGTGATCT
TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTCACGAAAGCCCCCTCTTCCCAGCCA
AGGAAGGAGTCTATGGGCCCCGCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC
CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG
CAGAGGGTGCCTTGCTGGCGGAGTCCCCCTCTTGGCTCTAGTGATGCATAGGGGAAACAGGGGA
CATGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACTTAC
TTGTGTAAC TGACAATTTCTGCAGAAATCCCCCTTCTCTAAATTCCCTTTACTCCACTGAG
GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA
TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCGGAGAA
TTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACA
TTAAAAAATAATTTCTACCAAATGGAAAGGAAATGTTCTATGTTGTTTCAGGCTAGGAGTAT
ATTGGTTCGAAATCCCAGGGAAAAAATAAAAAATAAAAAATTAAAGGATTGTTGAT

FIGURE 36

MRPAFALCLLWQALWPGPGGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS
TVRAGAE LRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE
SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLC PAPRPGA
ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTS GEGQPTLGGTGVPTRRPPATA
TSPVPQRTWP I RVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP
SGSVISKFNSTTSSATPQAFDSSSAVVFI FVSTAVVVLVILTMTVLGLVKLCFHESPSSQPR
KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 189-193, 381-385

Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 98-102, 434-438

Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,
477-483

Aspartic acid and asparagine hydroxylation site.

amino acids 262-274

Downloaded from www.jstor.org

FIGURE 37

CGGACGCGTGGGATTTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG
CGTCGAGTCAGACGGCACCATAATCGCCTTTAAAAGTGCCTCCGCCCTGCCGGCCGCGTATC
CCCCGGCTACCTGGGCGGCCCGCGGCGGTGCGCGCGTGAGAGGGAGCGCGCGGGCAGCCGA
GCGCCGCTGTGAGCCAGCGCTGCTGCCAGTGTGAGCGGCGGTGTGAGCGCGGTGGGTGCGGA
GGGGCGTGTGTGCCGGCGCGCGCGCCGTGGGGTGCAAACCCCGAGCGTCTACGCTGCCATGA
GGGGCGCGAACGCCTGGGCGCCACTCTGCCTGCTGCTGGCTGCCGCCACCCAGCTCTCGCGG
CAGCAGTCCCCAGAGAGACCTGTTTTTACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT
TATTGGCAGTGAAGGTTTTCTGGAGTGTACCCTCCAAATAGCAAATGTACTTGGAAAATCA
CAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTCGAGAGTGACAAC
CTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTTGGCCG
CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA
TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA
AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAAC
CCCCAACTGGCCAGACCGGGATTACCCTGCAGGAGTCACTTGTGTGTGGCACATTGTAGCCC
CAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC
CGATATGATTATGTGGCTGTGTTTAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA
GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAATGAACTTCTTATTCAGT
TTTTATCAGACTTAAAGTTTAACTGCAGATGGGTTTATTGGTCACTACATATTCAGGCCAAAA
AACTGCCTACAACACTACAGAACAGCCTGTCACCACCACATTCCCTGTAACCACGGGTTTTAA
ACCCACCGTGGCCTTGTGTCAACAAAAGTGTAGACGGACGGGGACTCTGGAGGGCAATTATT
GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAAACCATCACTCGCGATGGGAGTTTG
CACGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTTCAGCAGGCGGG
CAAGAACATGAGTGCCAGGCTGACTGTCGTCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTC
TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGC
TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG
TTAAACAGTGAACGTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTC
TCTCAGTAGAAAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCCGAAAAGATGG
GACTGGTTGACTCTTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAAGTTCTT
TGCTGCTGTCAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGA
AGCTAAAAGTGTCAAGCGTTGACAGCTTGGAAGCGTTTATTTATACATCTCTGTAAAAGGAT
ATTTTAGAATTGAGTTGTGTGAAGATGTCAAAAAAGATTTTAGAAGTGCAATATTTATAGT
GTTATTTGTTTACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTTCTA
AATCAATGCTTAATAAAATATTTTAAAGGAAAAAAAAAAAAA

CGGACGCGTGGGATTTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG

Figure 1 consists of 12 histograms arranged horizontally, labeled x_0 through x_{11} . Each histogram shows the frequency of values for x_k ranging from 0 to 10. The distributions are approximately normal, centered at 5. The peak frequency (count) for each x_k is as follows:

x_k	Peak Count
x_0	10
x_1	10
x_2	10
x_3	10
x_4	10
x_5	10
x_6	10
x_7	10
x_8	10
x_9	10
x_{10}	10
x_{11}	12

Signal sequence:

N-glycosylation site.

Casein kinase II phosphorylation site.

Tyrosine kinase phosphorylation site.

N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295,
305-311, 309-315, 320-326, 330-336

Cell attachment sequence.

amino acids 149-152

FIGURE 39

CGGACGCGTGGGCGGACGCGTGGGCGGCCCCACGGCGCCCGCGGGCTGGGGCGGTGCTTCTT
CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCATGGCCCCGAAGG
GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCTCAACCTCCCAGGACCTATCTGG
CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTG
CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACTTTGGAG
GTGGAAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTCAACAAGCAGCAGGAGGCCCCGGACCTCTTCC
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCGCAGGCACCTTCGGGCCCTCCTGC
CTTCCCTGTCCTGGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGG
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCCTGTGGCC
AGTGTGGCCTTGGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTTCGGCTTGT
TTTGGCCCCCTGTGCCCGATGCTCAGGACCTGAGGAATCAAACCTGTTTGCAATGCAAGAAGGG
CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAAC
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG
GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTGCTGTAAGAAGTGTAGCCCTGGCTATCA
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA
ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC
AGAAGACGAGTTGGTGGTGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA
CGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG
ACTGGCTACTGGTTGTCAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATATA
ATCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCAGAGCTTGGGCTGCCC
TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG
CCTTACAGAGCAGCCCAGGTACCCAGGCCCCGGGCAGACAAGGCCCTGGGGTAAAAAGTAGC
CCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGA
ATTTCAAAAGTTTTTCTTAATGGTGGCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTG
GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT
TCTGTGTTACCCACATCCCCACACCCCATTGCCACTTATTTATTCATCTCAGGAAATAAAGA
AAGGTCTTGGAAGTTAAAAA

[illegible]

Signal sequence:

Transmembrane domain:

N-glycosylation site.

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Casein kinase II phosphorylation site.

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

N-myristoylation site.

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179,
177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289,
326-332, 372-378, 395-401

Aspartic acid and asparagine hydroxylation site.

amino acids 321-333

EGF-like domain cysteine pattern signature.

amino acids 181-193

FIGURE 41

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA
GCACCATGCAGCCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCCCTGGCCAGCCCCGGG
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT
GCCCCACCCTGGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGGCCAGT
ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTTCAGCCAGAGC
TTCCGAGAGGTGGCCGGCAGGTTCTTGGCGTTGGAGGCCAGCACACACCTGCTGGTGTTCGG
CATGGAGCAGCGGCTGCCGCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGCGCCCCGGGCC
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA
ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGCTACAGGTGTCGGTGCAGAGG
GAGCATCTGGGCCCCGCTGGCGTCCGGCGCCCCACAAGCTGGTCCGCTTTGCCTCGCAGGGGGC
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTGGACCTTGGGGACTATGGAG
CTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG
ATGTACATTGACCTGCAGGGGATGAAGTGGGCGGAGAACTGGGTGCTGGAGCCCCCGGGCTT
CCTGGCTTATGAGTGTGTGGGCACCTGCCGGCAGCCCCCGGAGGCCCTGGCCTTCAAGTGGC
CGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA
GTGCAGCTGTGCCTCGGATGGTGCCTCGTGCCAAGGAGGCTCCAGCCATAGGCGCCTAGTG
TAGCCATCGAGGGACTTGACTTGTGTGTGTTTCTGAAGTGTTTCGAGGGTACCAGGAGAGCTG
GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTT
CCTCTGACAAGTTACCTCACCTAATTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA
GAGCCCTTGCTCAGTTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCANTGTGTCAATTGTTTACTTGTCTGTAC
TGGATCTGGGCTAAAGTCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT
TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAACATGAATAAAACACATTTTATTCT
AAAA

CGGAGGCTT

FIGURE 42

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV
ALLQRSHGDRSRGKRFSQSFREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP
VPKAALHRHGRLSPRSARARVTVEWLVRDGGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF
WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ
GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALVPRRLQP

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 76-80

Casein kinase II phosphorylation site.

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

Amidation site.

amino acids 74-78

TGF-beta family signature.

amino acids 282-298

U566650

[illegible]

GTCTGTTCCAGGAGTCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCGAATGGGGACAAA
GGCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTTACAGTGCACCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAACCTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCTACGAATCCCAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCTT
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGGTCTATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT
CTTGTTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTCTCTGGTGTGAGCCTGGTGGCTCACCGCCTATCATCTGCATTTGCCTTACT
CAGGTGCTACCGGACTCTGGCCCCCTGATGTCTGTAGTTTACAGGATGCCTTATTTGTCTTC
TACACCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC
ATCCTCCTTCATGCCCTCCCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACCTGTTTAAA
GTGTTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAAATGGCGGGGGTCGCAGGAATCTGCACTCAACTGCCCACCTGGC
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCTTGTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG
TGATGACACTGGGGTCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACCTGGAGGCTGGGCGCAGTGGCTCACGCCTG
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA
ACATGGAGAAACCTACTGGAAATACAAAGTTAGCCAGGCATGGTGTTGCATGCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

FIGURE 44

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLS CAYSGFSSPRVEW
KFDQGD TTRLVCYNNKITASYEDRVTF LPTGITFKSVTREDTGTYT CMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNI PSSATIGNRAVLTCSEQDGSPPSEYTWFKDGI VMPTNPKSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVA AVLVTLI
LLGILVFGIWFAYS RGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

Downloaded from www.sciencedirect.com

FIGURE 46

MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQAAGPSSSGSCPPTKQCR
TSGLCVPLTWRCRDLDCSDGSDEEEECRIEPCCTQKGQCPCPPPGGLPCPCTGVSDCSGGTDKKL
RNC SRLACLAGELRCTLSDDCIPLTWRC DGHPCDSSDELGCGTNEILPEGDATTMGPPVT
LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVTATLL
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

Signal sequence:

amino acids 1-30

Transmembrane domain:

amino acids 230-246

N-glycosylation site.

amino acids 126-130, 195-199, 213-217

Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation site.

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218,
224-230, 230-236, 263-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 44-55

Leucine zipper pattern.

amino acids 17-39

FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG
GTTAGACTGGCGGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG
GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG
GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA
TAATTCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG
TACCCCGACCTACACAATATGGTTTTATTATGTGCGGATGATGGAACGTGGAATAATCTGCC
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCGGTG
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCTGTCT
GCCACCCGCGGCCTTGTGAGCGCTACAACCACGGAACCTGTGGTGGAGTTTTACTGCGATCCT
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTTCTTC
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA
CCACGTGGAAGATTGTGGCGTTACGGCAACCAGTGTGCTGCTGGTGCTGCTGCTCGTCATC
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCAGGGGGCCTCCCCGGAG
TTCCAGCAGTGACCCTGACTTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG
ACGAAGCTGTGAGTGGCGGCTTGAGTGCTTAGGCCCGGGTACATGGCCTCTGTGGGCCAG
GGCTGCCCCCTTACCCGTGGACGACCAGAGCCCCCAGCATAACCCGGCTCAGGGGACACGGA
CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA
GTCTGTATTACCTCCCAGGTGCCAAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATA
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCCAGGCATCCATCATGCCCACTGGGT
GTTGTTCTTAAGAACTGATTGATTAAAAAATTTCCCAAAGTGTCCTGAAGTGTCTCTTCAA
ATACATGTTGATCTGTGGAGTTGATTCCTTTCTTCTTGGTTTTAGACAAATGTAAACAA
AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC
CTGTTTCTTCTTGACACAGACTGATTAAAAATTAAGNAAAAA

FIGURE 48

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGI
PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRI
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS
SNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLP
PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT
WPSTHETLLTTWKIVAFTATSVLLVLLLVLARMFQTKFKAHFPPRGPPRSSSSDPDFVVVD
GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGD TDTGPGESETCDS
VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEVASTSPGIHHAHWVLF LRN

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 325-344

N-glycosylation site.

amino acids 104-108, 134-138, 192-196

Casein kinase II phosphorylation site.

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,
364-368, 380-384, 467-471, 468-472

N-myristoylation site.

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,
478-484

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 394-405

Figure 1 displays 12 histograms, labeled x_0 through x_{11} , showing the distribution of the number of non-zero elements in the vector x_k for $k = 0, 1, \dots, 11$. The x-axis for each histogram represents the number of non-zero elements (0 to 10), and the y-axis represents the count (0 to 10). The distributions are roughly bell-shaped and centered around 5, with the peak count increasing from 10 at $k=0$ to 12 at $k=11$.

GCAGGAAAAAAAAA

FIGURE 50

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLLRGGQPVCRRGGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDFWIGLRRREEKQSNSTACQDL
YAWTDGSISQFRN WYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLLV
VTTVVCWVWICRKRKREQPD PSTKKQHTIWPSPHQNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES
GWVENEIYGY

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 296-300

Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,
299-303, 306-310, 323-327

N-myristoylation site.

amino acids 18-24, 37-43, 76-82, 146-152

0050040500

FIGURE 52

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMOVVFRHGAR
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL
TKVGMQQMFALGERLRKKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE
GPIIIHTDEADSEVLYPNYQSCWSLRQTRGRRTASLQPGISEDLKKVKDRMGIDSSDKVD
FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFHLHILES
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLMTLGIFDHKWPPFAVDLTMELYQHLESKEW
FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMS VYTLSP EKYHALCSQTQVMEVGNEE

Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 218-222

Casein kinase II phosphorylation site.

amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.

amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.

amino acids 50-65

U
S
P
A
T
E
N
T
O
F
F
I
C
E

FIGURE 53

CTCCTCTTAACATACTTGCAGCTAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT
TAAATTTAGCTCATCACCTTCACCTGCCTTGGTCATGGCTCTGCTATTCTCCTTGATCCTT
GCCATTTGCACCAGACCTGGATTCTTAGCGTCTCCATCTGGAGTGGGGCTGGTGGGGGGCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG
ATTGTTACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA
GTCCCAGAGGGTGTGAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA
GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGGCCGCAAAGGTGG
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA
GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTCG
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACCTGGGGAGAAAAGGAGGA
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTTCGTTGCTCAGGGGAGGAG
CAGTCCCTGGAGCAGTGCCAGCACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGA
TGTGGCTGTCATCTGCTCAGTGTAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCTGATTCTCAGGCCTTCAGAGTTGG
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA
CCACCTTTCCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACATA
CACCATTTGTCCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA
TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTTATTTACAATAATAAGATAGCAC
TATGTGTTCAAA

U5001101

[illegible]

Signal sequence:

Casein kinase II phosphorylation site.

N-myristoylation site.

Amidation site.

Speract receptor repeated domain signature.

amino acids 29-67, 249-287

FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG
GTGAAGGCCATGGACTTCATCACCTCCACAGCCATCCTGCCCCCTGCTGTTCCGGCTGCCTGGG
CGTCTTCGGCCTCTTCCGGCTGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT
GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC
GACATACTTGTCACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA
TGTGGACAAGAGGGTTCATGGAGACAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTGCGCCATCAGCAGCATCCAGGGCAAG
ATGAGCATTCCTTTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCCAGGCTTTCTTTGA
CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA
TCCACACCAACCTCTCTGTAAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC
ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGGCTGTTTATCTTCGAA
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC
AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC
TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAAGTGGG
AAAGACTGAAGAAACACATCTCGTGACAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA
AGCTTCTTCCCAGGGTGAGGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACACT
AAAAACTAGAAATAAACATCTCAAACAGTAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAG
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTAC

U
S
P
A
T
E
N
T
A
P
P
L
I
C
A
T
I
O
N

FIGURE 56

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA
KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL
VNNAGISYRGTIMDTTVDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAIISSIQKMSI
PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT
AQGRSPVEVAQDVLAAVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 104-120, 278-292

N-glycosylation site.

amino acids 228-232

Glycosaminoglycan attachment site.

amino acids 47-51

Casein kinase II phosphorylation site.

amino acids 135-139, 139-143, 253-257

Tyrosine kinase phosphorylation site.

amino acids 145-153, 146-153

N-myristoylation site.

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

Amidation site.

amino acids 265-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 6-17

03504031-071101
TOTAL "G" TAGS

FIGURE 57

CCCACGCGTCCGCTGGTGTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA
AAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAATTTCTTCTGGACATCCTC
CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCTCTAA
GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA
GACTGACTGCCTATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG
CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTTCATACCTTTGT
GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG
GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGCTACA
CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA
GGCATTTCCTTCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG
CTGGACATGTCTCGGTCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGCTGCTGTTGGA
TTTCATAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAACAACATG
TCTGTGTCCTAATTTGTAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTCTGACTGAGCAGAAGATG
ATTTTTATTCCATCTTCTATAGCTTTTTTAACAACATTGGAAAGGATCCTTCCTGAGCGTTT
CCTGGCAGTTTTTAAACGAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA
AAGCGCAATAAGCACCTAGTTTTCTGAAAAGTATTACCAGGTTTAGGTTGATGTCATCTA
ATAGTGCCAGAATTTTAATGTTTGAAGTCTGTTTTTTCTAATTATCCCCATTTCTTCAATA
TCATTTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT
ACATATGATATAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAAGAA
CCAAAATGACTTTATTAAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA
AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTGGATTGCACTTAAATTTTGT
ATAATTTGTGTTTCTTTTTCTGTTCTACATAAAATCAGAACTTCAAGCTCTCTAAATAAAA
TGAAGGACTATATCTAGTGGTATTTACAATGAATATCATGAACTCTCAATGGGTAGGTTTC
ATCCTACCCATTGCCACTCTGTTTCCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT
GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG
GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAA
AA
AA

U.S. GOVERNMENT PRINTING OFFICE

FIGURE 58

MKFLLDILLLLPLLIIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK
LVLWDINKHGLEETAACKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAIEIGDVSILVNNAGVV
YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC
SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVNRLMH
GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34, 283-287

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

Downloaded from www.sciencedirect.com

[illegible]

AAAATATTATATATAAAAGTAAAAAAAAAA

FIGURE 60

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDSGQPLFLTPYIEAGKIQKGRELSL
VGPFPGNLNPKSYAGFLTVNKTYNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH
GPYVVTSMNLTDRDFPWTTLTSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF
FQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYSDPESIIGG
YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVGTG
CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT
EIMNMYKVLINQGLDIIVAAALTELSLMGMDWKGSQEYKKAEEKVWKIFKSDSEVAGYIRQ
AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 81-85, 132-136, 307-311, 346-350

Casein kinase II phosphorylation site.

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,
353-357, 424-428

Tyrosine kinase phosphorylation site.

amino acids 423-432

N-myristoylation site.

amino acids 22-28, 110-116, 156-162, 232-238

Serine carboxypeptidases, serine active site.

amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

FIGURE 61

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACAT
TTTTCCCTTTCCTAACAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT
GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTG
CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG
TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC
CTGGGGGAGGGCCTGCCTAACAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT
AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGG
TGAACTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA
TAGAGATGCTTTGTAAAATAAAATTTTAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA
GACCAAAATAGATAACAGGATTCCTGAACATTCTAAGAGGGAGAAAGTATGTTAAAAATA
GAAAAACCAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCTGGGTC
AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTGGTCTGACCACTCTGCCTTGTTGTTTT
GCAGAATCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT
CCTCACCGCCGCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG
ACAGCCGCTCTGTGGTCTGCTCTCAGTGGTCTGGGTGCTGCTGGCCCCCCCCAGCAGCCGGC
ATGCCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT
CCACCAAGGGACGGGGGCCGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA
ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG
CCCCTCATCGTGCAGCCCTGCAGCGAAGTGCTCACCCCTACCAACAATGTCAACAAGCTGCT
CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA
AGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTGGAGCCATCCACAAGAAGGAGCACTAC
CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA
TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA
GCCGGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGAT
TTTGTCTCCTCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT
CTTCTACATCTACGGCTTTTGCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCGAGA
CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG
CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC
CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC
AGGCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAG
CAGTATCACCACCCGCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTT
GCAGATCAAGGAGCGCCTGCAGTCCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC
TGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCCATCGATGATAACTTCTGTGGA
CTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC
CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG
TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC
ATTACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTAACTATAG
GCAACTTTATTTTCTTGGGGAACAAAGGTGAAATGGGGAGGTAAGAAGGGGTAAATTTTGTG
ACTTAGCTTCTAGCTACTTCCCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA
TTTCAATATTTCCCAAACCTTTAAGAAAAAATTTAAGAAGGTACATCTGCAAAAGCAA

FIGURE 62

MGTLGQASLFAPPGNYFWSDHSALCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP
RALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGPEEDNKSRYPPPLIVQPCSEVLTLTNNVNKLLIIDYSENRLACGSL
YQGVCKLLRLDDLFI LVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY
FPTLSSRKLP RDPRESSAMLDYELHSD FVSSLIKIPSDTLALVSHFDIFYIYG FASGGFVYFL
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSL AQAFNITSQDDVLFAIFSKGQKQYHHPD DSALCAFP IRAINLQIKERLQSCYQGEGN
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG
YSVVFVGTKSGKLKKVRVYEFRC SNAIHLLSKESLLEGSYWWRFN YRQLYFLGEQR

Signal sequence:

amino acids 1-32

Transmembrane domain:

amino acids 71-87

N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,
384-388, 471-475, 481-485, 530-534

N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

FIGURE 63

AGGCTCCCGCGCGCGGCTGAGTGCAGGACTGGAGTGGGAACCCGGGTCCCGCGCTTAGAGAACACGCGATGACCA
CGTGAGCCTCCGGCGGAGGCCGGCCCGCACGCTGGGACTCCTGCTGCTGGTCGTCTTGGGCTTCTGGTGCTCC
GCAGGCTGGACTGGAGCACCTTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGA
ACTTCATGCTGGAGGATTCCACCTTCTGGATCTTGGGGGCTCCATCCACTATTTCCGTGTGCCAGGGAGTACT
GGAGGGACCGCCTGCTGAAGATGAAGCCCTGTGGCTTGAACACCTCACCACCTATGTTCCGTGGAACCTGCATG
AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCTTCGTCTGATGGCCCGCAGAGATCG
GGCTGTGGGTGATTCTGCGTCCAGGCCCCCTACATCTGCAGTGAGATGGACCTCGGGGGCTTGCCAGCTGGCTAC
TCCAAGACCCTGGCATGAGGCTGAGGACAACCTTACAAGGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACC
TGATGTCCAGGGTGGTGCCACTCCAGTACAAGCGTGGGGGACCTATCATTGCCGTGCAGGTGGAGAATGAATATG
GTTCTCTATAATAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAATGC
TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGATTGTCCAGGGAGTCTTGGCCACCATCAACTTGCAGT
CAACACACGAGCTGCAGCTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGATGGTGAATGCAGT
ACTGGACGGGGTGGTTTACTCGTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT
CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATG
GAGCCATGCACTTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG
ATTACACGGCCAAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCACCTG
ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTTGTACCTGTCTCTGTGGGACGCCCTCAAGTACC
TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCCCT
TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCCTCAGTGGCCACGTGCATGATCGGGGGCAGG
TGTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT
ACACCGTGCTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATTGATGACCAGCGCAAAG
GCTTAATTGGAATCTCTATCTGAATGATTCACCCCTGAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA
GCTTCTTTAGAGGTTCCGCTTGACAAATGGNGTTCCTCCCAGAAACACCCACATTACCTGCTTTCTTCTTGG
GTAGCTTGTCCATCAGCTCCACGCCCTTGTGACACCTTTCTGAAGCTGGAGGCTGGGAGAAGGGGGTTGTATTCA
TCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCAGAAGACGCTTTACCTCCCAGGTCCCTGGTTGA
GCAGCGGAATCAACCAGGTATCGTTTTTGGAGAGACGATGGCGGGCCCTGCATTACAGTTCACGGAAACCCCCC
ACCTGGGCAGGAACCAGTACATTAAGTGAAGCGGTGGCACCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCTC
CTCTTGACCTGAAGCCTGGTGGCTGCTGCCCCACCCCTCACTGCAAAGCATCTCCTTAAGTAGCAACCTCAGGG
ACTGGGGGCTACAGTCTGCCCCGTCTCAGCTCAAACCCCTAAGCCTGCAGGGAAAGGTGGGATGGCTCTGGGCC
TGGCTTTGTTGATGATGGCTTTCTACAGCCCTGCTCTTGTGCCGAGGCTGTCCGGCTGTCTTAGGGTGGGAGC
AGCTAATCAGATCGCCAGCCTTTGGCCCTCAGAAAAAGTGCTGAAACGTGCCCTTGACCCGGACGTCACAGCCC
TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTTGCTGGTTCTTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT
TTTATCCCCGAAATCCTGGGTGTGTACCAAGTGTAGAGGGTGGGGAAGGGGTGTCTACCTGAGCTGACTTTGTT
CTTCCTTCAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACCTCGGCGTGAGAAACATGTGACTTCCCTT
TCCCTTCCCACTCGCTGCTTCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCTGCGTCTTCC
CAAGTTAGCAGGTGTCTCTGGTGTTTCAAGTGGAGGAGCATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCA
CATCCAGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCC
AGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGA
GGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG
ACAGAAGGCCAGCTCAGTGGCCCCCGCTCCCCACCCCCACGCCCCAAGCAGGGGGCAGAGCAGCCCTCCTTC
GAAGTGTGTCAAGTCCGCATTTGAGCCTTGTCTGGGGCCAGCCCAACACCTGGCTTGGGCTCACTGTCCTGA
GTTGCAGTAAAGCTATAACCTTGAATCAAA

U
S
P
A
T
E
N
T
A
P
P
L
I
C
A
T
I
O
N

FIGURE 64

MTTWSLRRRPPARTLGLLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW
IFGGSIHVFRVPREYWRDRLLKMKACGLNTLTITYVPWNLHEPERGKFDFSGNLDLEAFVLMA
AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQ
YKRGGPPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLTSDNKGLSKGIVQGVLAT
INLQSTHELQLLTTFLFNVQGTQPKMVMETWTGWFDWSGGPHNILDSSEVLKTVSAIVDAGS
SINLYMFHGGTNFGFMNGAMHFHDKSDVTSYDYDAVLTEAGDYTAHYMKLRDFFGSI SGIP
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYE
TSITSSGILSGHVHVRGQVFVNTVSI GFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN
IDDQRKGLIGNLYLNDSPKKNFRIYSLDMKKSFFQRFGLDKWXSLEPETPTLPAFFLGSL SIS
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSGGINQVIVFEETMAGPA
LQFTETPHLGRNQYIK

Signal sequence:

amino acids 1-27

Casein kinase II phosphorylation site.

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

N-myristoylation site.

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

FIGURE 64

FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC
CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG
GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCATGGCTCCC
AAGAAGCTGTCCTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCA
GGCAGACACTCGGTTCGTTCTAGTGGATAGGGGTTCATGACCGGTTTCTCCTAGACGGGGCCC
CGTTCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC
CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACCTA
CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCCTTCTGA
ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG
TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC
AGATCCAGACTTCCTTGCCGCAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATC
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATTAGGTGGAGAATGAATATGGTAGC
TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACTGCTAGG
AGAAAAGATCTTGCTCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGG
GACTCTATAACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA
CTGGGGGCCAGAATCACTCCACACGGTCTGTGTGAGCTGTAACCAAAGGACTAGAGAACATGC
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG
AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC
TATATCTGAAGCAGGGGACCCACACCTAAGCTTTTTGCTCTTCGAGATGTCATCAGCAAGT
TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCCGAGCCCCAAGATGATGCTTGGACCTGTG
ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCAT
TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC
GAACCTATATGACCCATAACCTTTTGGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC
CATGACCGTGCCATATGTGATGGTGGATGGGGTGTTCAGGGTGTGTGGAGCGAAATATGAG
AGACAAACTATTTTTGACGGGGAAACTGGGGTCCAACTGGATATCTTGGTGGAGAACATGG
GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG
GGGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG
GTGGTTTCCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT
ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG
ACCAAGGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCCGGTACTGGACAAAGCAGGGGCC
ACAACAGACCCTCTACGTGCCAAGATTCTGCTGTTTCTAGGGGAGCCCTCAACAAAATTA
CATTGCTGGAAC TAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTGGATAAGCCTATC
CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACTAG
TGCCTCTGAACCAATGGAGTTAAGTGGGCACTGAAAGGTAGGCCGGGCATGGTGGCTCATGC
CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTGAGGACTTCAAGA
CCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG
ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC
AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA
GACACTCCATCTCAAAAAAAAAAAAA

FIGURE 66

MAPKKLSCLRSLLLPLSLTLLLPQADTRS FVVDGRGHDRFLLDGAPFRYVSGSLHYFRVPRVL
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLA AVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFS YMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF
TLLRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNM MYFHGGTNF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPPSPKMML
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVPN
NGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGDTFLYL
PGWTKGQVWINGFNLGRYWTKQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

Signal sequence:

amino acids 1-27

N-glycosylation site.

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 4-8

Casein kinase II phosphorylation site.

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,
603-607, 644-648

Tyrosine kinase phosphorylation site.

amino acids 191-198

N-myristoylation site.

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

FIGURE 67

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC
ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT
CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT
GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT
CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG
AAAACTCAGGCAGCACATTTACGCAACGCCAGGACAAGCAGGAGTTGCATCTGTTTCATG
CTGTGCGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAAC
AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC
TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA
TGCCTTCACGTGAAGTTCACTGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA
CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTACATAATGACGGCAC
TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACCTCCAGA
ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT
TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT
GACTTGTTTTAAATTTATGGCATAACAAAATTGTTACTATTCTCCCTCTATTACCCATGTCA
AAAACCTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT
AGTTTACAGAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA
AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC
CAAAACAATTGTTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC
TCACTCCCAGAGAAAGTTGGTCAGCTCTCCAGCTCACTCAGCTGGAGCTGAAGGGGAACTG
CTTGGAACCGCTGCCAGCCCAGCTGGGCCAGTGTGCGATGCTCAAGAAAAGCGGGCTTGTTG
TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA
AATATTCCCTTTGCAAATGGGATTTTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC
AATTCCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG
ATACATCTTTTAAAATAAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT
GTTCAATGTTTGTAGGGTTTTAAGTCATTCATTTCCAAATCATTTTTTTTTTTCTTTTGGGG
AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTTCTTTTTTAAATTGTTTGTAACCTGGAT
GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT
TTCTTACTAAAAAAAAAAAAAAAAA

[illegible]

Signal sequence:

N-glycosylation site.

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Casein kinase II phosphorylation site.

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

Tyrosine kinase phosphorylation site.

amino acids 349-355, 375-381

N-myristoylation site.

amino acids 78-84, 124-130, 212-218, 392-398

FIGURE 69

CCCACGCGTCCGGCCTTCTCTCTGGACTTTGCATTTCCATTCCCTTTTCATTGACAAACTGACTTTTTTTTATTTCT
TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTTACACACATAAGGAT
CTGTGTTTTGGGGTTTCTTCTTCTCCCTGACATTGGCATTGCTTAGTGTTGTGTGGGGAGGGAGACCACGTGG
GCTCAGTGCTTGCCTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC
ATCGCTGGTGGTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTGCTGCTCTGTCTTTACTTCAAAATACACAAC
GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAAGGTGTGGTGGGCCAAG
AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGCTCCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC
AGTTTTGATTCCCTGCCACCTTGCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAA
GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG
CTCCCCATCAGTTTCATGGAAAATAACTCAGTGCCCTGCTGGGAACCACTGCTGGAGATCCCTACAGAGAGCTTC
CACTGGGGGCAACCCTTCCAGGAAGGAGTTGGGGAGAGAGAACCCTCACTGTGGGGAATGCTGATAAACCACTCA
CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGCTGGAACCTGACGTTTCCCTGGAGGTGTCCAGAAA
GCTGATGTAACACAGAGCCTATAAAAGCTGTGCGTCTTAAGGCTGCCAGCGCCTTGCCAAAATGGAGCTTGTA
AGAAGGCTCATGCCATTGACCCTCTTAATTCTCTCCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT
GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC
AGTGAGAATGCACCTGGACAATAGAAAGACCAGAAAACAAAGCATCAGAATTATCTTTTCCATGTCCAGCTT
GATCCAGATGGAAGCTGTGAAAGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCTCTGCTAGGG
CAAGTCTGCAGTAAAAACGACTATGTTCCCTGTATTTTGAATCATCATCCAGTACATTGACGTTTCAAATAGTTACT
GACTCAGCAAGAATTCAAAGAACTGTCTTTGTCTTCTACTACTTCTTCTCTCTAACATCTCTATTCCAACTGT
GGCGGTTACCTGGATACCTTGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT
TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAACTAAACTTCAAAGAGATTTTCCCTAGAAATAGAC
AAACAGTGCAAATTTGATTTTCTTGCCATCTATGATGGCCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT
GGCCGTGTGACTCCCACCTTCGAATCGTCATCAAACCTCTCTGACTGTGCTGTTGTCTACAGATTATGCCAATTCT
TACCGGGGATTTTCTGCTTCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT
GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA
GACCCAACTTGACAGACCAAATTTATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA
AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC
CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA
GAAGATGATGTAATACAAAGTCAAAATGCACCTGGGCAAAATATAACACCAGCATGGCTCTTTTGAATCCAATTCA
TTTGAAGAGCTATACTTGAATCACCATATTATGTGGATTTGAACCAAACCTTTTTTGTTCAGTTAGTCTGCAC
ACCTCAGATCCAAATTTGGTGGTGTCTTCTGATACCTGTAGAGCCTCTCCACCTCTGACTTTGCATCTCCAACC
TACGACCTAATCAAGAGTGGATGTAGTGCAGATGAACTTGTAAAGGTGTATCCCTTATTTGGACACTATGGGAGA
TTCCAGTTTAATGCCTTTAAATTCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAGTTTTGATATGTGAT
AGCAGTGACCACCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG
AAAACAGATTCCATCATAGGACCCATTCTGTCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAGGATTTTCAGCAT
GAAACACATGCGGAAGAACTCCAAACCAGCCTTTCAACAGTGTGCATCTGTTTTCTTTCATGGTTCTAGCTCTG
AATGTGGTGAAGTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG
CAGAACTATTAACCTAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT
GGCTACACATATTATGAATAAATGAGGAAGGGCTGAAAGTGACACACAGGCCTGCATGTAAAAAAA

Signal sequence:

Transmembrane domain:

N-glycosylation site.

Casein kinase II phosphorylation site.

Tyrosine kinase phosphorylation site.

N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 71

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG
GGACATGCGGCCCCAGGAGCTCCCCAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGCTGCTG
TGCTGCCGCGCCGCGCCGTGCCCTGCCACAGCGCCACGCGCTTCGACCCACCTGGGAGTCC
CTGGACGCCCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG
GGGAGTGTTTTCCGTGCCAGCTTCGGTAGCGAGTGGTCTGGTGGTATTGGCAAAAGGAAA
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTTCAAATATGAAGAT
TTTGGAACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTTCAGGC
CTCTGGTGCCAAATACATTGTCTTAACCTTCCAACATCATGAAGGCTTTACCTTGTGGGGGT
CAGAATATTCGTGGAACCTGGAATGCCATAGATGAGGGGCCAAGAGGGACATTGTCAAGGAA
CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTGGACTGTACTATTCCCTTTTTTGA
ATGGTTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACATATCAGCCTGAGGTTCTGTGG
TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGCCCTGGTT
ATATAATGAAAGCCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA
TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGCATTGTTATAACCCAGGACATCTTTTGCCA
CATAAATGGGAAAACCTGCATGACAATAGACAACTGTCTCTGGGGCTATAGGAGGGAAGCTGG
AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG
GAGGAAATCTTTTGATGAATATTGGGGCCACACTAGATGGCACCATTCTGTAGTTTTTGGAG
GAGCGACTGAGGCAAGTGGGGTCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA
TACCTGGCGATCCAGAAATGACACTGTCAACCCAGATGTGTGGTACACATCCAAGCCTAAAG
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCACATCAGGACAGCTGTTCTTGGCCAT
CCCAAAGCTATTCTGGGGGCAACAGAGGTGAACTACTGGGCCATGGACAGCCACTTAACTG
GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATGC
CGTGTAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG
ATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA
AAGCAATGTAACTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTTCCCTTTTTCCACTA
AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACCTTTGCCATTAAAGTC
TCTTCACATTGATTTGTTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTTCACATTATAGTAG
CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG
ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT
TTTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA
CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG
AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG
GTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG
CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTA
TTGAACACTGCCAGACGTTATAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAA
GTTTTTCTTTCTTCAATTATAAATAACATAAGTGTACTGTAACTTTACAAACGTTTTTAATT
TTTAAACCTTTTTGGCTCTTTTGTAAATAACACTTAGCTTAAACATAAACTCATTGTGCAA
ATGTAA

FIGURE 72

MRPQELPRLAFPLLLLLLLLLLLLLPPPPCPAHSATRFDPWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEFWFWYWQKEKIPKYVEFMKDNYPSPSKYEDFGPLFTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLEW
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKWGWALALTNVI

Signal sequence:

amino acids 1-28

N-glycosylation site.

amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319,
375-375

Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.

amino acids 410-432

Alpha-L-fucosidase putative active site.

amino acids 283-295

10960403040506070809101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

FIGURE 73

U990401.071

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC
TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTG
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT
TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCTTGTGAGCAAAAAGGCGAACCAGC
AGCTGAATTTACAGAAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGGAGA
TGGATTTCGTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG
TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGAT
ACTTGGAATAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA
AACTGCAACACAAACAACAGAATTTATTGTCAGTGACAGTACCTACTCGGTGGCATCCCCCTT
ACTCTACAATACCTGCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG
AGAAAAAAATTGATTTGTGTACAGAAAGTTTTTATGGAACTAGCACCATGTCTACAGAAAC
TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC
CCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGGATTTTGC
TATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAAATGAT
CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA
AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCCTGGAA
GCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTTCATGCTCC
TTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAAAGAAGAAAGTCCACCCTT
GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC
CCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAGCTTCCCACGGCC
TTTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAA
GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG
TGGGTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGAC
CCTTCTTTCAGCTCTGAAAGAGAAACACGTATCCACCTGACATGTCCTTCTGAGCCCGGTA
AGAGCAAAAAGAAATGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG
ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT
GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA
ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT
AGGAAATATACTTTTACAAGTAACAAAAATAAAACTCTTATAAATTTCTATTTTTATCTGA
GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAGTAATAAAATTCA
ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT
TGAATATTATTCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTTTCA
GTTTTGATATTTCTAGCTTATCTACTTCCAAACTAATTTTTTATTTTTGCTGAGACTAATCTT
ATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG
TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAAACAAATGTATCACTA
GCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA
AGCATTTAGAAAACCTT

FIGURE 74

MARCFSLVLLLLTSIWTTLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR
LLGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQF
AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP
APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNAAAGFGGVPTALLVLALLFF
GAAAGLGFCYVKRYVKAFFFTNKNQOKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP
SKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

Downloaded from www.jstor.org

FIGURE 75

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTTCGGTGCCGCGACTTTCACGATGG
CTCGCCCAACCTTACTACCTTCTGTGCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA
ACTGCCGCCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTG
ACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA
TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT
TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC
TGATGACGTGCAAACCCCCCTATATATGGGCCCCTGAGTATATCAAGTACTTCAATGATAAA
ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA
TTGGTCTAATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAATACAAC
GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC
AAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG
AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAACTATCA
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTGAGA
TGGGGAAAACAAGAAGGATAAATAAGATCCTCACTTTGGCAGTGCTTCCTCTCCTGTCAATT
CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG
CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG
GCACCCTACAGGAAGGCCTGCCATGCTGTGGCCAACTGTTTCACTGGAGCAAGAAAGAGATC
TCATAGGACGGAGGGGAAATGGTTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC
AGCTATTCAGACATCTCCATGGTTTTCTCCATGAAACTCTGTGGTTTTCATCATTCCTTCTTAG
TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA
CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC
TTCATTAAAAGTATAAGCCTAACTTTGTGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG
TTTTTATCATTGAAGACAATATTGAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGAA
TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC
TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCGTAATAAAAAGATTGGGATT
TCCTTTTG

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
0	0	1	4	9	16	25	36	49	64	81	100	121	144	169	196	225	256	289	324	361	400	441	484	529	576	625	676	729	784	841	900	961	1024	1089	1156	1225	1296	1369	1444	1521	1600	1681	1764	1849	1936	2025	2116	2209	2304	2401	2500	2601	2704	2809	2916	3025	3136	3249	3364	3481	3600	3721	3844	3969	4096	4225	4356	4489	4624	4761	4900	5041	5184	5329	5476	5625	5776	5929	6084	6241	6400	6561	6724	6889	7056	7225	7396	7569	7744	7921	8100	8281	8464	8649	8836	9025	9216	9409	9604	9801	10000

amino acids 1-48

amino acids 111-125

amino acids 165-169, 185-189

amino acids 154-158, 265-269

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

amino acids 188-194, 225-231

amino acids 244-253

FIGURE 77

GGACAGCTCGCGGCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG
GGGCCCCAGCCTGGCCCCGGGTCAACCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA
TTGCTCCTGCTGCCCGGCTCCTACGGA CTGCCCTTCTACAACGGCTTCTACTACTCCAACAG
CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG
TGGTGGAGACACCCGAGGAGACCCTGTTACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC
CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAAGCT
GTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT
TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG
ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCATTGACGGGCTGGA
GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCCTTACCAGTCCCCCA
ACGGGCGCTACCAGTTCAACTTCCACGAGGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG
GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGAGGGCCTGGACTGGTGC AACGCGGG
CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCGGCAGCCCTGCGGTGGCC
CAGGCCTGGCACCTGGCGTGC GAAGCTACGGCCCCCGCCACCGCCGCTGCACCGCTATGAT
GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCCCTGAGAAGCT
GACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC
AGCTCTTTGCCGCTGGAAGTTCCATGGCCTGGACCGCTGCGACGCTGGCTGGCTGGCAGAT
GGCAGCGTCCGCTACCCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT
CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTG TACGGTGTTTACTGCTACCGCCAGC
ACTAGGACCTGGGGCCCTCCCCTGCCGCATTCCCTCACTGGCTGTGTATTTATTGAGTGGTT
CGTTTTCCCTTGTGGGTTGGAGCCATTTTAACTGTTTTTATACTTCTCAATTTAAATTTTCT
TTAAACATTTTTTTTACTATTTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG
GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTGTGGGCTTCTG
GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTCCCTTAAGGAGGTTGGTGGCCAGAGTGGGC
GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCT
CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCTTTGGGCCTCTCCTGCC
CACCTCTACTTCTCTGTGAAGCCGCTGACCCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA
GCCAGTTCTAGGCTTCCAGGCGAAATCTGAGGGAAGGAAGAACTCCCCTCCCCGTTCCCCT
TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTGTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG
GGAGGGGCCCTCAGGTGTGTGTACTTTGGACAATAAATGGTGTCTATGACTGCCTTCCGCCAA
AAA
AAA

11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 841. 842. 843. 844. 845. 846.

Signal sequence:

Casein kinase II phosphorylation site.

Tyrosine kinase phosphorylation site.

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

FIGURE 79

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC
TTCTGTTGCTACTGAGGCACGGGGCCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC
CAGGGGAGGGTGCACCAGGCGGCCCCCTGAGCGACGCTCCCCATGATGACGCCCACGGGAA
CTTCCAGTACGACCATGAGGCTTTCTGAGCGGGAAGTGGCCAAGGAATTCGACCAACTCA
CCCCAGAGGAAAGCCAGGCCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC
GGCGACGGCTGGGTGTCGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA
CATACGGGACTCGGTGAGCGCGGCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGG
GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCCGGTGAAGAATTTTCATGAC
GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGC
CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCCTGCACCCCGAGG
AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAA
GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA
GGAGCCGGCGTGGGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGG
ATGGGCACCTGGATGGGAGTGAAGTGGGCCACTGGGTGCTGCCCCCTGCCCAGGACCAGCCC
CTGGTGGAAAGCCAACCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA
AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG
ACCTGACCCGGCACCCACGATGAGCTGTGAGCACCGCGCACCTGCCACAGCCTCAGAGGCCCG
CACAATGACCGGAGGAGGGGGCCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCCGCGAGGAG
GCAGATGCAGTCCCAGGCATCCTCCTGCCCCCTGGGCTCTCAGGGACCCCCTGGGTCTGGCTTC
TGTCCTCTGTCACACCCCCCAACCCAGGGAGGGGCTGTATAGTCCCAGAGGATAAGCAATAC
CTATTTCTGACTGAGTCTCCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAA
GAACCGCCCCAACCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCT
GGCCCCAGCCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAA
AAGCCAGCGCCGGGACCTTGAAA
AAAAAAAAAAAAAAAAA

U
S
P
A
T
E
N
T
A
P
P
L
I
C
A
T
I
O
N

| | | | | | | | | | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1987 | 1986 | 1985 | 1984 | 1983 | 1982 | 1981 | 1980 | 1979 | 1978 | 1977 | 1976 | 1975 | 1974 | 1973 | 1972 | 1971 | 1970 |
| 1969 | 1968 | 1967 | 1966 | 1965 | 1964 | 1963 | 1962 | 1961 | 1960 | 1959 | 1958 | 1957 | 1956 | 1955 | 1954 | 1953 | 1952 |
| 1951 | 1950 | 1949 | 1948 | 1947 | 1946 | 1945 | 1944 | 1943 | 1942 | 1941 | 1940 | 1939 | 1938 | 1937 | 1936 | 1935 | 1934 |
| 1933 | 1932 | 1931 | 1930 | 1929 | 1928 | 1927 | 1926 | 1925 | 1924 | 1923 | 1922 | 1921 | 1920 | 1919 | 1918 | 1917 | 1916 |
| 1915 | 1914 | 1913 | 1912 | 1911 | 1910 | 1909 | 1908 | 1907 | 1906 | 1905 | 1904 | 1903 | 1902 | 1901 | 1900 | 1899 | 1898 |
| 1897 | 1896 | 1895 | 1894 | 1893 | 1892 | 1891 | 1890 | 1889 | 1888 | 1887 | 1886 | 1885 | 1884 | 1883 | 1882 | 1881 | 1880 |
| 1879 | 1878 | 1877 | 1876 | 1875 | 1874 | 1873 | 1872 | 1871 | 1870 | 1869 | 1868 | 1867 | 1866 | 1865 | 1864 | 1863 | 1862 |
| 1861 | 1860 | 1859 | 1858 | 1857 | 1856 | 1855 | 1854 | 1853 | 1852 | 1851 | 1850 | 1849 | 1848 | 1847 | 1846 | 1845 | 1844 |
| 1843 | 1842 | 1841 | 1840 | 1839 | 1838 | 1837 | 1836 | 1835 | 1834 | 1833 | 1832 | 1831 | 1830 | 1829 | 1828 | 1827 | 1826 |
| 1825 | 1824 | 1823 | 1822 | 1821 | 1820 | 1819 | 1818 | 1817 | 1816 | 1815 | 1814 | 1813 | 1812 | 1811 | 1810 | 1809 | 1808 |
| 1807 | 1806 | 1805 | 1804 | 1803 | 1802 | 1801 | 1800 | 1799 | 1798 | 1797 | 1796 | 1795 | 1794 | 1793 | 1792 | 1791 | 1790 |
| 1789 | 1788 | 1787 | 1786 | 1785 | 1784 | 1783 | 1782 | 1781 | 1780 | 1779 | 1778 | 1777 | 1776 | 1775 | 1774 | 1773 | 1772 |
| 1771 | 1770 | 1769 | 1768 | 1767 | 1766 | 1765 | 1764 | 1763 | 1762 | 1761 | 1760 | 1759 | 1758 | 1757 | 1756 | 1755 | 1754 |
| 1753 | 1752 | 1751 | 1750 | 1749 | 1748 | 1747 | 1746 | 1745 | 1744 | 1743 | 1742 | 1741 | 1740 | 1739 | 1738 | 1737 | 1736 |
| 1735 | 1734 | 1733 | 1732 | 1731 | 1730 | 1729 | 1728 | 1727 | 1726 | 1725 | 1724 | 1723 | 1722 | 1721 | 1720 | 1719 | 1718 |
| 1717 | 1716 | 1715 | 1714 | 1713 | 1712 | 1711 | 1710 | 1709 | 1708 | 1707 | 1706 | 1705 | 1704 | 1703 | 1702 | 1701 | 1700 |
| 1699 | 1698 | 1697 | 1696 | 1695 | 1694 | 1693 | 1692 | 1691 | 1690 | 1689 | 1688 | 1687 | 1686 | 1685 | 1684 | 1683 | 1682 |
| 1681 | 1680 | 1679 | 1678 | 1677 | 1676 | 1675 | 1674 | 1673 | 1672 | 1671 | 1670 | 1669 | 1668 | 1667 | 1666 | 1665 | 1664 |
| 1663 | 1662 | 1661 | 1660 | 1659 | 1658 | 1657 | 1656 | 1655 | 1654 | 1653 | 1652 | 1651 | 1650 | 1649 | 1648 | 1647 | 1646 |
| 1645 | 1644 | 1643 | 1642 | 1641 | 1640 | 1639 | 1638 | 1637 | 1636 | 1635 | 1634 | 1633 | 1632 | 1631 | 1630 | 1629 | 1628 |
| 1627 | 1626 | 1625 | 1624 | 1623 | 1622 | 1621 | 1620 | 1619 | 1618 | 1617 | 1616 | 1615 | 1614 | 1613 | 1612 | 1611 | 1610 |
| 1609 | 1608 | 1607 | 1606 | 1605 | 1604 | 1603 | 1602 | 1601 | 1600 | 1599 | 1598 | 1597 | 1596 | 1595 | 1594 | 1593 | 1592 |
| 1591 | 1590 | 1589 | 1588 | 1587 | 1586 | 1585 | 1584 | 1583 | 1582 | 1581 | 1580 | 1579 | 1578 | 1577 | 1576 | 1575 | 1574 |
| 1573 | 1572 | 1571 | 1570 | 1569 | 1568 | 1567 | 1566 | 1565 | 1564 | 1563 | 1562 | 1561 | 1560 | 1559 | 1558 | 1557 | 1556 |
| 1555 | 1554 | 1553 | 1552 | 1551 | 1550 | 1549 | 1548 | 1547 | | | | | | | | | |

Signal sequence:

N-glycosylation site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Endoplasmic reticulum targeting sequence.

amino acids 325-330

FIGURE 81

GGGGCCTTGCCCTTCGCACTCGGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG
GCGGCGGGGCGCGGGTGCAGGGGATCCCTGACGCCCTCTGTCCCTGTTTCTTTGTGCTCCCAG
CCTGTCTGTGCTCGTTCGTTTTTGGCGCCCCCGCCTCCCCGCGGTGCGGGGTTCACACCGATCCTG
GGCTTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCCTGGAGCAG
CGGGTCGTCTGTGTCCTCTCTCCTCTGCGCCGCGCCCCGGGGATCCGAAGGGTGCAGGGGCTCT
GAGGAGGTGACGCGCGGGGCCTCCCGCACCTTGGCCTTGCCCGCATTCTCCCTCTCTCCCAG
GTGTGAGCAGCCTATCAGTCACCATGTCGCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG
TGTCTGCTGCTGCTGCCGGGGCCCGGGGAGCGAGGGAGCCGCTCCCATTGCTATCACATG
TTTTACCAGAGGCTTGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCCAGGGGGCTGCC
CTCTTGAGGAATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG
GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC
TGGTCGAGAAAACATTTCCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT
GGTCTGCTTCTTTTACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA
GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACCTGG
CAATAAAGATTGTAAAGCAGACATTGCATTTCTGATTGATGGAAGCTTTAATATTGGGCAGC
GCCGATTTAATTTACAGAAGAATTTTGTGGGAAAGTGGCTCTAATGTTGGGAATTGGAACA
GAAGGACCACATGTGGGCCTTGTTCAGCCAGTGAACATCCCAAATAGAAATTTACTTGAA
AACTTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTTCAGAGGGGTA
ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGA
GTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTATTGATGGTTGGCCTTCTGATGACAT
CGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA
AGCCTATCCCTGAAGAACTGGGGATGGTTCCAGGATGTACATTTGTTGACAAGGCTGTCTGT
CGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAA
GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT
CAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTGAGATAGCAATTTCCGCCTC
ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT
AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTTCACTGACTATAGCACCA
AAGAGAATGTCTAGCTGTCATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT
GATGCCATTTCTTCACTGTTAGAAATGTGTTTGGCCCTATAAGGGAGAGCCCCAACAAGAA
CTTCTAGTAATTGTACAGATGGGCAGTCCATGATGATGTCCAAGGCCCTGCAGCTGCTG
CACATGATGCAGGAATCACTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG
AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTTCTTCAAGAGAGTTTACAGGATT
AGAACCAATTGTTTCTGATGTCATCAGAGGCATTTGTAGAGATTTCTTAGAATCCCAGCAAT
AATGGTAACATTTTGACAACTGAAAGAAAAAGTACAAGGGGATCCAGTGTGTAAATTGTATT
CTCATAATACTGAAATGCTTTAGCATACTAGAATCAGATACAAAACCTATTAAGTATGTCAAC
AGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT
ACTTTGTTAAAAACACTGCTGAGGCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA
GATAATGTGGATTAAACCTTAAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAA
TTCCATAGCTCAATAAAAGAATCTGATACTTAGACCAAAAAAAAAA

FIGURE 82

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPPGGCPLEEF SVY
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASF TVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDC KADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLNFTSAKDVLFAI KEVGFGRGNSNTGKAL
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG
MVQDVTTFVDKAVCRNNGFFSYHMPNWF GTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI
DGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAHDAGITI
FSVGVAWAPLDDLKDMASKPKESHAF FTREFTGLEPIVSDVIRGICRDFLESQQ

Signal sequence:

amino acids 1-24

N-glycosylation site.

amino acids 100-104, 221-225

Casein kinase II phosphorylation site.

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,
425-429, 478-482, 528-532

N-myristoylation site.

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,
239-245, 240-246, 261-267, 403-409, 442-448, 443-449

Amidation site.

amino acids 145-149

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCAGCCC
GGCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGGCCGAGCGCAACTCGGTCCAGTCG
GGGCGGCGGCTGCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGC
TGCTGGCGGCGGCGGTCCCCACGGCCCCCGGCCGCTCCGACGGCGACCTCGGCTCCAGTC
AAGCCCCGGCCGGCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGA
GGTTGAGGAAGCTGATGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGG
CAGAAGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTAT
CACAATGAGACCAACACAGACACGAAGTTTGAAATAATACCATCCATGTGCACCGAGAAAT
TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTG
TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGC
ATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCT
CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAA
TGGCCACCAGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG
TGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA
GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG
GAGCCTTGGAACGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG
GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC
CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG
AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCT
GCACTGCTGGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA
ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACA
TCTTCTTCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGC
TCCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAAC
TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG
ACAGCCGTTTGTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGG
AGTCTCCCTCTGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAAACATCAA
CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCCTTCCATGGGCATAGGTAAGCTG
TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCC
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC
CCTCTCTCAGCACAGCCTGGGGAGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCACACCAGCCTTGGTGCCACCAA
AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTCTTGAGGCATGCACATCTGGAATTAAG
GTCAAACATAATTCTCACATCCCTCTAAAAGTAAACTACTGTTAGGAACAGCAGTGTTCTCAC
AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT
TGCATTAGTAACCTTGAAAGGTATATGACTGAGCGTAGCATAACAGGTTAACCTGCAGAAACA
GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC
TGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT
GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTTCAGGTGTCA
TGGACTGTTGCCACCATGTATTCATCCAGAGTTCCTTAAAGTTTAAAGTTGCACATGATTGTA
TAAGCATGCTTTCTTTGAGTTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC
ATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

[illegible]

Signal sequence:

N-glycosylation site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Amidation site.

amino acids 140-144

FIGURE 85

AAGGAGGCTGGGAGGAAAGAGGTAAGAAAGGTTAGAGAACCTACCTCACATCTCTCTGGGCTCAGAAGGACTCTG
AAGATAACAATAATTTTCAGCCCATCCACTCTCCTTCCCTCCCAAACACACATGTGCATGTACACACACACATACA
CACACATACACCTTCTCTCTTCACTGAAGACTCACAGTCACTCACTCTGTGAGCAGGTCATAGAAAAGGACAC
TAAAGCCTTAAGGACAGGCTGGCCATTACCTCTGCAGCTCCTTTGGCTTGTGAGTCAAAAAACATGGGAGGGG
CCAGGCACGGTGACTCACACCTGTAATCCAGCATTGTTGGGAGACCGAGGTGAGCAGATCACTTGAGGTGAGGAG
TTCGAGACCAGCCTGGCCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCAGGAGTGTTGGC
AGGTGCCTGTAATCCAGCTACTCAGGTGGCTGAGCCAGGAGAATCGCTTGAATCCAGGAGGCGGAGGATGCAGT
CAGCTGAGTGACCGCTGCCTCCAGCCTGGGTGACAGAATGAGACTCTGTCTCAAACAAACAAACACGGGAGGA
GGGGTAGATACTGCTTCTCTGCAACCTCCTTAACCTCTGCATCCTCTTCTTCCAGGGCTGCCCTGATGGGGCCTG
GCAATGACTGAGCAGGCCCAGCCCCAGAGGACAAGGAAGAGAAGGCATATTGAGGAGGGCAAGAAGTGACGCCCG
GTGTAGAATGACTGCCCTGGGAGGGTGGTTCCCTTGGGCCCTGGCAGGGTTGCTGACCCCTACCCCTGCAAAACACA
AAGAGCAGGACTCCAGACTCTCCTTGTGAATGGTCCCCTGCCCTGCAGCTCCACCATGAGGCTTCTCGTGGCCCC
ACTCTTGCTAGCTTGGGTGGCTGGTGCCACTGCCACTGTGCCCGTGGTACCCTGGCATGTTCCCTGCCCCCTCA
GTGTGCCTGCCAGATCCGGCCCTGGTATACGCCCGCTCGTCTACCGCGAGGCTACCCTGTGGACTGCAATGA
CCTATTCTGACGGCAGTCCCCCGGCACTCCCCGAGGCACACAGACCCCTGCTCCTGCAGAGCAACAGCATTGT
CCGTGTGGACCAGAGTGAGCTGGGCTACCTGGCCAATCTCACAGAGCTGGACCTGTCCAGAACAGCTTTTCGGA
TGCCCGAGACTGTGATTTCCATGCCCTGCCCCAGCTGCTGAGCCTGCACCTAGAGGAGAACCAGCTGACCCGGCT
GGAGGACCACAGCTTTGACAGGGCTGGCCAGCCTACAGGAACCTATCTCAACCACAACCAGCTCTACCGCATCGC
CCCCAGGGCCTTTTCTGGCCTCAGCAACTTGCTGCGGCTGCACCTCAACTCCAACCTCCTGAGGGCCATTGACAG
CCGCTGGTTTGAATGCTGCCCAACTTGGAGATACTCATGATTGGCGGCAACAAGGTAGATGCCATCCTGGACAT
GAACTTCCGGCCCTGGCCAACCTGCGTAGCCTGGTGTAGCAGGCATGAACCTGCGGGAGATCTCCGACTATGC
CCTGGAGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACAGCTGGCCCGGGTGCCAGGCGGGCACT
GGAACAGGTGCCCGGGCTCAAGTTCCTAGACCTCAACAAGAACCCGCTCCAGCGGGTAGGGCCGGGGGACTTTGTC
CAACATGCTGCACCTTAAGGAGCTGGGACTGAACAACATGGAGGAGCTGGTCTCCATCGACAAGTTTGCCCTGGT
GAACCTCCCCGAGCTGACCAAGCTGGACATCACCAATAACCCACGGCTGTCTTCATCCACCCCCGCGCCTTCCA
CCACCTGCCCCAGATGGAGACCCCTCATGCTCAACAACAACGCTCTCAGTGCCTTGACCCAGCAGACGGTGGAGTC
CCTGCCCAACCTGCAGGAGGTAGGTCTCCACGGCAACCCCATCCGCTGTGACTGTGTATCCGCTGGGCCAATGC
CACGGGCACCCGTGTCCGCTTCATCGAGCCGCAATCCACCCTGTGTGCGGAGCCTCCGGACCTCCAGCGCCTCCC
GGTCCGTGAGGTGCCCTTCCGGGAGATGACGGACCACTGTTTGGCCCTCATCTCCCCACGAAGCTTCCCCCAAG
CCTCCAGGTAGCCAGTGGAGAGAGCATGGTGTGCTGCAATGCGGGCACTGGCCGAACCCGAACCCGAGATCTACTG
GGTCACTCCAGCTGGGCTTCGACTGACACCTGCCCATGCAGGCAGGAGGTACCGGGTGTACCCCGAGGGGACCCT
GGAGCTGCGGAGGGTGACAGCAGAAGAGGCAGGGCTATACACCTGTGTGGCCAGAACCTGGTGGGGGCTGACAC
TAAGACGGTTAGTGTGGTTGTGGGCCGTGCTCTCCTCCAGCCAGGCAGGGACGAAGGACAGGGGCTGGAGCTCCG
GGTGCAGGAGACCCACCCCTATCACATCCTGCTATCTTGGGTACCCCAACCAACACAGTGTCCACCAACCTCAC
CTGGTCCAGTGCCTCCTCCCTCCGGGGCCAGGGGGCCACAGCTCTGGCCCGCTGCCTCGGGGAACCCACAGCTA
CAACATTACCCGCCTCCTTCAGGCCACGGAGTACTGGGCCTGCCTGCAAGTGGCCTTTGCTGATGCCACACCCA
GTTGGCTTGTGTATGGGCCAGGACCAAGAGGCCACTTCTTGCCACAGAGCCTTAGGGGATCGTCTTGGGCTCAT
TGCCATCCTGGCTCTCGCTGTCTTCTCCTGGCAGCTGGGCTAGCGGCCACCTTGGCACAGGCCAACCAGGAA
GGGTGTGGGTGGGAGGCGGCCTCTCCCTCCAGCCTGGGCTTCTTGGGGCTGGAGTGCCCCCTTCTGTCCGGGTGT
GTCTGCTCCCCTCGTCTGCTGCCCTGGAATCCAGGGAGGAAGCTGCCAGATCCTCAGAAGGGGAGACACTGTTGCC
ACCATTTGTCTCAAAATTCTTGAAGCTCAGCCTGTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTTTACCAA
AGAGAAGCAGTCTGGGCCAGATGCCCTGCCAGGAAAGGGACATGGACCCACGTGCTTGAGGCCCTGGCAGCTGGGC
CAAGACAGATGGGGCTTTGTGGCCCTGGGGGTGCTTCTGCAGCCTTGAAAAAGTTGCCCTTACCTCCTAGGGTCA
CCTCTGCTGCCATTCTGAGGAACATCTCCAAGGAACAGGAGGGACTTTGGCTAGAGCCTCCTGCCTCCCCATCTT
CTCTCTGCCAGAGGCTCCTGGGCCTGGCTTGGCTGTCCCTACCTGTGTCCCCGGGCTGCACCCCTTCTCTTCT
TCTTTCTCTGTACAGTCTCAGTTGCTTGTCTTGTGCCCTCCTGGGCAAGGGCTGAAGGAGGCCACTCCATCTCAC
CTCGGGGGGCTGCCCTCAATGTGGGAGTGACCCAGCCAGATCTGAAGGACATTTGGGAGAGGGATGCCAGGAA
CGCCTCATCTCAGCAGCCTGGGCTCGGCATTCCGAAGCTGACTTTCTATAGGCAATTTTGTACCTTTGTGGAGAA
ATGTGTACCTCCCCCAACCCGATTCACTCTTTTCTCCTGTTTTGTAAAAAATAAAAAATAATAATAACAATAAA
AAAA

11990401.1 = 07.1.01.1

FIGURE 86

MRLLVAPLLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA
VPPALPAGTQTL LLQSNSIVRVDQSELGYLANLT ELDLSQNSFSDARDCDFHALPQLLSLHL
EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE
MLPNLEILMIGGNKVDAILDMNFRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQ
LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELV SIDKFALVNLP
ELTKLDITNNPRLSFIHPRAFHHL PQMETLMLNNA LSALHQQTVESLPNLQEVGLHGNPIR
CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPPFREMTDHCLPLISPRSFPPSLQ
VASGESMVLHCRALAEPEPEIYWVTPAGLR LTPAHAGRRYRVYPEGTLELRRVTAE EAGLYT
CVAQNLVGADTKTVSVVVGRALLQPGRDEGQGLELRVQETHPYHILLSWVTPPNTVSTNL TW
SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS
CHRALGDRPGLIAILALAVLLLAAGLAHLGTGQPRKGVGGRRPLPPAWAFWGWSAPSVRV V
SAPLVLPWNPNRKLPRSSGETLLPPLSQNS

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 629-648

N-glycosylation site.

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 485-489

Casein kinase II phosphorylation site.

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146,
243-247, 313-317, 488-492, 700-704

Tyrosine kinase phosphorylation site.

amino acids 532-540

N-myristoylation site.

amino acids 15-21, 493-499, 566-572

Amidation site.

amino acids 470-474, 660-664, 692-696

00990401-07101

[illegible]

ATCTATAACAGAAAAAAAAAAAAAAAAA

| | 1970 | 1971 | 1972 | 1973 | 1974 | 1975 | 1976 | 1977 | 1978 | 1979 | 1980 | 1981 | 1982 | 1983 | 1984 | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 | 2026 | 2027 | 2028 | 2029 | 2030 | 2031 | 2032 | 2033 | 2034 | 2035 | 2036 | 2037 | 2038 | 2039 | 2040 | 2041 | 2042 | 2043 | 2044 | 2045 | 2046 | 2047 | 2048 | 2049 | 2050 | 2051 | 2052 | 2053 | 2054 | 2055 | 2056 | 2057 | 2058 | 2059 | 2060 | 2061 | 2062 | 2063 | 2064 | 2065 | 2066 | 2067 | 2068 | 2069 | 2070 | 2071 | 2072 | 2073 | 2074 | 2075 | 2076 | 2077 | 2078 | 2079 | 2080 | 2081 | 2082 | 2083 | 2084 | 2085 | 2086 | 2087 | 2088 | 2089 | 2090 | 2091 | 2092 | 2093 | 2094 | 2095 | 2096 | 2097 | 2098 | 2099 | 2100 | 2101 | 2102 | 2103 | 2104 | 2105 | 2106 | 2107 | 2108 | 2109 | 2110 | 2111 | 2112 | 2113 | 2114 | 2115 | 2116 | 2117 | 2118 | 2119 | 2120 | 2121 | 2122 | 2123 | 2124 | 2125 | 2126 | 2127 | 2128 | 2129 | 2130 | 2131 | 2132 | 2133 | 2134 | 2135 | 2136 | 2137 | 2138 | 2139 | 2140 | 2141 | 2142 | 2143 | 2144 | 2145 | 2146 | 2147 | 2148 | 2149 | 2150 | 2151 | 2152 | 2153 | 2154 | 2155 | 2156 | 2157 | 2158 | 2159 | 2160 | 2161 | 2162 | 2163 | 2164 | 2165 | 2166 | 2167 | 2168 | 2169 | 2170 | 2171 | 2172 | 2173 | 2174 | 2175 | 2176 | 2177 | 2178 | 2179 | 2180 | 2181 | 2182 | 2183 | 2184 | 2185 | 2186 | 2187 | 2188 | 2189 | 2190 | 2191 | 2192 | 2193 | 2194 | 2195 | 2196 | 2197 | 2198 | 2199 | 2200 | 2201 | 2202 | 2203 | 2204 | 2205 | 2206 | 2207 | 2208 | 2209 | 2210 | 2211 | 2212 | 2213 | 2214 | 2215 | 2216 | 2217 | 2218 | 2219 | 2220 | 2221 | 2222 | 2223 | 2224 | 2225 | 2226 | 2227 | 2228 | 2229 | 2230 | 2231 | 2232 | 2233 | 2234 | 2235 | 2236 | 2237 | 2238 | 2239 | 2240 | 2241 | 2242 | 2243 | 2244 | 2245 | 2246 | 2247 | 2248 | 2249 | 2250 | 2251 | 2252 | 2253 | 2254 | 2255 | 2256 | 2257 | 2258 | 2259 | 2260 | 2261 | 2262 | 2263 | 2264 | 2265 | 2266 | 2267 | 2268 | 2269 | 2270 | 2271 | 2272 | 2273 | 2274 | 2275 | 2276 | 2277 | 2278 | 2279 | 2280 | 2281 | 2282 | 2283 | 2284 | 2285 | 2286 | 2287 | 2288 | 2289 | 2290 | 2291 | 2292 | 2293 | 2294 | 2295 | 2296 | 2297 | 2298 | 2299 | 2300 | 2301 | 2302 | 2303 | 2304 | 2305 | 2306 | 2307 | 2308 | 2309 | 2310 | 2311 | 2312 | 2313 | 2314 | 2315 | 2316 | 2317 | 2318 | 2319 | 2320 | 2321 | 2322 | 2323 | 2324 | 2325 | 2326 | 2327 | 2328 | 2329 | 2330 | 2331 | 2332 | 2333 | 2334 | 2335 | 2336 | 2337 | 2338 | 2339 | 2340 | 2341 | 2342 | 2343 | 2344 | 2345 | 2346 | 2347 | 2348 | 2349 | 2350 | 2351 | 2352 | 2353 | 2354 | 2355 | 2356 | 2357 | 2358 | 2359 | 2360 | 2361 | 2362 | 2363 | 2364 | 2365 | 2366 | 2367 | 2368 | 2369 | 2370 | 2371 | 2372 | 2373 | 2374 | 2375 | 2376 | 2377 | 2378 | 2379 | 2380 | 2381 | 2382 | 2383 | 2384 | 2385 | 2386 | 2387 | 2388 | 2389 | 2390 | 2391 | 2392 | 2393 | 2394 | 2395 | 2396 | 2397 | 2398 | 2399 | 2400 | 2401 | 2402 | 2403 | 2404 | 2405 | 2406 | 2407 | 2408 | 2409 | 2410 | 2411 | 2412 | 2413 | 2414 | 2415 | 2416 | 2417 | 2418 | 2419 | 2420 | 2421 | 2422 | 2 |
|--|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|---|
|--|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|---|

Transmembrane domain:

N-glycosylation site.

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Casein kinase II phosphorylation site.

N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447

FIGURE 89

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCATGGAGCTGGCACTGCGGCGCTCTCCCGT
CCCGCGGTGGTTGCTGCTGCTGCCGTGCTGCTGGGCCTGAACGCAGGAGCTGTCATTGACT
GGCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG
TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCTGGTCAT
GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTTGAAACTTTGAGGAAATTGGGC
CCCTTGACAGTGATCTCAAACCACGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT
GTGGATAATCCCGTGGGCACTGGGTTTCACTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA
CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG
AATTCCAGACAGTTCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC
ATTGGTCTAGAGCTTTATAAGGCCATTTCAGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT
TGCCTTGGGTGATTCTGATCTCCCCTGTTGATTTCGGTGCTCTCCTGGGGACCTTACCTGT
ACAGCATGTCTCTTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA
CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT
GATCATTGAACAGAACACAGATGGGGTGAAGTTCTATAACATCTTAAGTAAAGCACTCCCA
CGTCTACAATGGAGTCGAGTCTAGAATTCACACAGAGCCACCTAGTTTGTCTTTGTCAGCGC
CACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAA
GCTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA
TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC
AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATAACCATGGGTGAGGAGGCCTG
GGTGCGGAAACTGAAGTGGCCAGAACTGCCTAAATTCAGTCAGCTGAAGTGGAAAGGCCCTGT
ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC
TACTGGATTCTGAAAGCTGGTCATATGGTTTCTTGACCAAGGGGACATGGCTCTGAAGAT
GATGAGACTGGTGAAGTCTCAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT
TGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT
AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT
GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAACCTAAGATTTTTTAAAAAATTGAT
TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

FIGURE 90

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC
KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSDLKPRKTTWLQAASLLFVDNPVGTGFSY
VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR
GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVUNKGLYRE
ATELWGKAEMII EQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQORDALS
QLMNGPIRKKLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL
IVDTMGQEAWVRKLKWPPELPKFSQLKWKALYSDPKSLETSFVKS YKNLAFYWILKAGHMVP
SDQGDMAKMMRLVTQQE

Signal sequence:

amino acids 1-25

N-glycosylation site.

amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 101-105

Casein kinase II phosphorylation site.

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-myristoylation site.

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,
187-193, 195-201, 331-337, 332-338, 360-366

FIGURE 91

GGCCGCGGGAGAGGAGGCCATGGGCGCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTC
GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCGGCGCCGTTATCAGGACCATGCGGCCGA
CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCCGAACCTCGGGCGTTGGCCGTGGCA
GGGGAGCCTGCGCCTGTGGGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG
CACTCACGGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG
GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG
TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCCTATGACATTG
CCTTGGTGAAGCTGTCTGCACCTGTCACCTAAACACATCCAGCCCATCTGTCTCCAG
GCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA
AGAGGATGAGGCACTGCCATCTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAACA
ACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG
GTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT
GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG
GTCGGCCCAATCGGCCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAAG
CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTTCCCTCT
TCTCTGGGCTCTCCCACTCCTGGGGCCGGTCTTGAGCCTACCTGAGCCCATGCAGCCTGGGGC
CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTGGTAATAAACACATTCCAGTTGA
TGCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA

U590001.1

FIGURE 92

MGARGALLLALLLARAGLRKPESQEAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLW
DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI
YLSPLYLGNSPYDIALVKLSAPVITYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP
SPHTLQEVQVAIINNSMCNHLFLKYSFRKDIFGDMVCAGNAQGGKDACFGDSSGGPLACNKNG
LWYQIGVVSWSGVGCRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL
LGPV

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 167-171, 200-204, 273-277

Casein kinase II phosphorylation site.

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

N-myristoylation site.

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245,
259-265, 269-275

Amidation site.

amino acids 33-37

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 252-263,

Serine proteases, trypsin family, histidine active site.

amino acids 78-84

FIGURE 93

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGAATGGGACTCCAAGCCTGCCTCCTAGGGCT
CTTTGCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGAGCCCGACCAGCGGAGGACGC
TGCCCCCAGGCTGGGTGTCCCTGGGCCGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT
GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTGCGGATCCCAG
CTCTCCTCAATACGGAAAATACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC
CACTGACCCTCCACACGGTGCAAAAATGGCTCTTGCGAGCCGGAGCCAGAAAGTGCCATTCT
GTGATCACACAGGACTTTTCTGACTTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC
TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCAC
ATCCCTACCAGCTTCCACAGGCCTTGGCCCCCATGTGGACTTTGTGGGGGGACTGCACCGT
TTTCCCCAACATCATCCCTGAGGCAACGTCCTGAGCCGAGGTGACAGGGACTGTAGGCCT
GCATCTGGGGGTAAACCCCTCTGTGATCCGTAAGCGATACAACCTTGACCTCACAAGACGTGG
GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCAGTTCTTGAGCAGTATTTTCATGAC
TCAGACCTGGCTCAGTTCATGCGCCTCTTCGGTGGCAACTTTCACATCAGGCATCAGTAGC
CCGTGTGGTTGGACAACAGGGGCCGGGGCCGGGCGGGATTGAGGCCAGTCTAGATGTGCAGT
ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG
GGACAGGAGCCCTTCTGCACTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT
GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA
AACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCTGCTCTTCGCCTCAGGTGACAGT
GGGGCCGGGTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG
CCCTATGTCACCACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCAAAATGAAA
TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCACGGCCTTCATACCAGGAG
GAAGCTGTAACGAAGTTCCTGAGCTCTAGCCCCACCTGCCACCATCCAGTTACTTCAATGC
CAGTGGCCGTGCCTACCCAGATGTGGCTGCACTTTCTGATGGCTACTGGGTGGTGAGCAACA
GAGTGCCCATTTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTGGGGGGATCCTA
TCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCTCTTGGCTTTCTCAACCCAAG
GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC
TGGATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCCTGGCTGGGATCCTGTAACAGGC
TGGGGAACACCAACTTCCCAGCTTTGCTGAAGACTCTACTCAACCCCTGACCCTTTCCTATC
AGGAGAGATGGCTTGTCCCTGCCCCGAAGCTGGCAGTTCAGTCCCTTATTCTGCCCTGTTG
GAAGCCCTGCTGAACCCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAA
TGCTGTGAGCTTGACTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT
CCTGCCTTAGATTCTCAATAAGATGCTGTAACCTAGCATTTTTTGAATGCCTCTCCCTCCGC
ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCAAAGGGTTGTATACAGACTCTGTGCACTA
TTTCACTTGATATTCAATCCCCAATTCAGTGCAGGAGACCTCTACTGTCACCGTTTACTCT
TTCCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATG
GCCTTTCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACCTCTCTG
ACTACTCTTGTCTTCTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC
TCCATTTGTAGATTTTTGCTCTTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACA
TCTACAACCATTACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAA
TGTAATAAA

U950401.101

FIGURE 94

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL
SIRQAELLLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGGLHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGLFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

Signal sequence:

amino acids 1-16

N-glycosylation site.

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

Glycosaminoglycan attachment site.

amino acids 361-365, 408-412, 538-542

Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225,
248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488,
521-527, 533-539, 549-555

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted October 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 10

[illegible]

FIGURE 96

MAGIPGLLFLLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIIYLSSSGDGAQHRDSGSSGKS
RRKRQIYGYSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG
TQKL RVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA
LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD
AQP GASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW
IKGNYLDCREG

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 93-97, 207-211

Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

050401-0101
T0101-0101

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCCGCCCCACCAGCCATGGTGGTTT
CTGGAGCGCCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG
GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA
GCAGCTGAACCGGGTGTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA
GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC
ACTGCTGCCCCTGTGTTTCAAGGACAACCTGAACAAACCATACTGTTCTCTGTGCTGCTGGG
GGCCTGGCAGCTGGGGAACCTGGCTCTCGGTCCCAGAAGGTGGGTGTTGCCTGGGTGGAGC
CCCACCCTGTGTATTCTTGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCCTCTCGAG
CGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC
CCCACCCTCAGACCCTGCAGAAGCTGAAGGTTCTATCATCGACTCGGAAGTCTGCAGCCAT
CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT
GGAGGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGGCCCCCTCATGTGCCAGGTGGACG
GCGCCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCC
GGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA
GCTCCGCGGGGCGGCTCAGGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG
CCGCGCGCTCCTTAGGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACA
TCTGGATCTGGATCTGCGGCGGCCTCGGGCGGTTTCCCCCGCCGTAAATAGGCTCATCTACC
TCTACCTCTGGGGGCCCCGGACGGCTGCTGCGGAAAGGAAACCCCCTCCCCGACCCGCCCGAC
GGCCTCAGGCCCCCTCCAAGGCATCAGGCCCCGCCCAACGGCCTCATGTCCCCGCCCCAC
GACTTCCGGCCCCCGCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT
AGGTATTTGTAACCCTGCCACATATCTTATTTATTCCTCCAATTTCAATAAATTATTTATT
CTCCAAAAAAAAA

[illegible]

```
><subunit 1 of 1, 317 aa, 1 stop
```

MVVSGAPPALGGGCLGTFSTSLLLLASTAILNAARIPVPPACGKQPQLNRVVGGEDSTDSEWP
 WIVSIQKNGTHHCAGSLTSTRWVITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA
 WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPNTHCWISGWGSIQDG
 VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMC
 QVDGAWLLAGIISWGECAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG
 SGAAARS

amino acids 1-32

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

amino acids 826-830

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

amino acids 607-615

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

THE UNIVERSITY OF CHICAGO

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCTCTGATGCTTCTGCTGCCGCTACTGCTA
CTGCTGGTGGCCACCACAGGCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGTATGGT
GGAGCTGCACAACCTCTACCGGGCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA
GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCCTGTGGGGCCAC
AACAAGGAGCGCGGGCGCCGCGGCGAGAATCTGTTCCGCATCACAGACGAGGGCATGGACGT
GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT
GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT
GGTGTGCAACTATGAGCCTCCGGGGAAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC
CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC
CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGA
AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT
TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAACCCAGGCC
CCAACCTCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC
AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA
CCTTCCCCAAATCGACCCATGTTCCCTATCCCAAATCAGCAGACAAAGTGACAGACAAAACA
AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG
GGAACCTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG
AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC
CACACGGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAA
TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCGTCTTGCCAGGTGCAGAGGGCCCTGACA
AGCCTAGCGTTGTGTGTCAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGA
CTACTGCTCCTGCCTCCTCTGGTGTTGGCTGGAATCTTCTGAATGGGATACCACTCAAAGGG
TGAAGAGGTCAGCTGTCTCCTGTCTATCTTCCCCACCCTGTCCCAGCCCCCTAAACAAGATA
CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC
ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCGCTGAGGACTGCACACC
GGGCCCACACCTCTCCTGCCCCCTCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCT
CACTGCCTACCTGGCCTGGGGCTGTCTGCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG
TGCTAGCTGGGGATGGGGATTCTTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC
TTTGAGTGGGGGAGGCAGGGACGAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCT
GTCCAACCTGTGAAA

FIGURE 100

MHGSCSFLMLLLPLLLLLVATTGPVGALTDEEKRLMVELHNLRYAQVSPTASDMLHMRWDEE
LAFAKAYARQCVWGHNKERGRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ
MCGHYTQVVWAKTERIGCGSHFCEKLQGVETNIELLVCNYEPPGNVKGKRPYQEGTPCSQC
PSGYHCKNSLCEPIGSPEDAQDLPLYLVTEAPSFRAEASDSRKMGTTPSSLATGIPAFLVTEV
SGSLATKALPAVETQAPTSLATKDPSPMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKS
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEAEELPPSSEVLAS
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV
SGLNSGPGHVWGPLLGLLLLLPPLVLAGIF

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237,
250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

FIGURE 101

GTAAGTGAAGTCAGGCTTTTCATTGGGAAGCCCCCTCAACAGAATTGCGTCATTCTCCAAGTTATGGTGGACGT
ACTTCTGTTGTTCTCCCTCTGCTTGTCTTTTTCACATTAGCAGACCGGACTTAAGTCACAACAGATTATCTTTTCAT
CAAGGCAAGTTCCATGAGCCACCTTCAAAGCCTTCGAGAAGTGAACTGAACAACAATGAATTGGAGACCATTCC
AAATCTGGGACCAGTCTCGGCAAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGA
ACATCTGAAAGAGTTTCAGTCCCTTGAACTTTGGACCTTAGCAGCAACAATATTTTCAAGCTCCAAACTGCATT
TCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAA
TTTGGCCAACACACTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAACT
GCCCCAACTGCAACATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGG
TGCTCTGAAGTCTCTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTGGGGGCTGAGCAA
CATGGAAATTTTGCAGCTGGACCATAACCAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGTGATGCT
GCAGGAACCTCATCTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAG
TGAGCTGGACCTAACTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCCTAAGCTTACTAAATAC
ACTGCACATTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCTTCCGGGGGCTTTCCAGTTTAAAGACTTT
GGATCTGAAGAACAATGAAATTTCTGGACTATTGAAGACATGAATGGTGTCTTCTCTGGGCTTGACAAACTGAG
GCGACTGATACTCCAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCA
TCTAGACCTGAGTGACAACGCAATCATGCTTTTACAAGGCAATGCATTTTACAAATGAAGAACTGCAACAATT
GCATTTAAATACATCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCACAGTGGGTGGCGGAAAACAACCTT
TCAGAGCTTTGTAATGCCAGTTGTGCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTGTCTGTTAGCCGAGA
TGGCTTTGTGTGTGATGATTTTCCCAAACCCAGATCACGGTTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTT
CAATTTGAGTTTTCATCTGCTCAGCTGCCAGCAGCAGTGATTCCCAATGACTTTTGTCTGGAAAAAAGACAATGA
ACTACTGCATGATGCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCAC
CATCCTTCCGGCTGCGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCTCATCTCCAATCACTTTGGTTT
ATCCTACTCTGTCAAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTACCAAGACCCCCATGGATCTCACCAT
CCGAGCTGGGGCCATGGCAGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCAGATAGCCTGGCAGAAGGA
TGGGGGCACAGACTTCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCGAGGATGACGTGTTCTTTATCGT
GGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTTCAAGCAATGC
AACTCTGACTGTCTTAGAAACACCATCATTTTGTGGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGC
CGTCTTACAGTGCAATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAGAGATGATAGCCCATTTGGTGGT
AACCAGAGAGGCATTTTGTGAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTGATGCTGGGAA
ATACAGATGTGAGATGCTAACACCCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAAC
CTGCGACTCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGGGCACTGTGGGTGTCGTGATCATAGC
CGTGGTTTGTGTGTGGTGGGCACGTCACTCGTGTGGGTGGTTCATCATATACCACACAAGGCGGAGGAATGAAGA
TTGCAGCATTACCAACACAGATGAGACCAACTTGCCAGCAGATATTCTAGTTATTTGTCTCATCTCAGGGAACGTT
AGCTGACAGGCAGGATGGGTACGTGTCTCAGAAAGTGGAAGCCACCACAGTTTGTACATCTTTCAGGTGCTGG
ATTTTCTTACCACAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGC
CACAGATCTGTTCCCTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGA
TCCTTTTGAACATATCATACAGGTTGCAGTCTTGACCCAAGAACAGTTTAAATGGACCACTATGAGCCAGTTA
CATAAAGAAAAAGGAGTGCTACCCATGTTCTCATCCTTCAAGAAGATCCTGCGAACGGAGCTTCAGTAATATATC
GTGGCCTTTCATGTGAGGAAGCTACTTAACACTAGTTACTCTACAATGAAGGACCTGGAATGAAAAATCTGTG
TCTAAACAAGTCCTCTTTAGATTTTAGTGCAATCCAGAGCCAGCGTCGGTTGCCCTCGAGTAATCTTTTCATGGG
TACCTTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTGAGCC
AAGAGCCTTTTATTTGAAAGCTCATTCTTCCCGAGACTTGGACTCTGGGTGAGAGGAAGATGGGAAAGAAAGGAC
AGATTTTTCAGGAAGAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACCTACAGGACTCCAAATTTTCAGTC
TTATGACTTGGACACATAGACTGAATGAGACCAAGGAAAAGCTTAAACATACTACCTCAAGTGAACCTTTTATTTA
AAAGAGAGAGAATCTTATGTTTTTTAAATGGAGTTATGAATTTTAAAAGGATAAAAAATGCTTTATTTATACAGAT
GAACCAAAATTACAAAAGTTATGAAAATTTTATACCTGGGAATGATGCTCATATAAGAATACCTTTTTTAAACTA
TTTTTTAACTTTGTTTTATGCAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTT
TTATAATGCCAGATTTCTTTTTATGGAAGATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTT
TTAAATAGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAA

U.S. Patent & Trademark Office

FIGURE 102

MVDVLLLFSLCLLFHISRPLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVVSAN
ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPPALQLKYLYLNSNRVTSMEP
GYFDNLANTLLVLKLNRRNRI SAIPPKMFKL PQHQHLELNRNKIKNVDGLTFQGLGALKSLKM
QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWE
FCQKLSELDLTFNHL SRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSSLKTLDLKNNE
ISWTIEDMNGAFSGLDKLRRLLIQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ
MKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFSQSFVNASCAHPQLLKGRSIFAVSPDGFVCD
DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKDNELLHDAEMENY AHLRAQG
GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA
MARLECAAVGHPAPQIAWQKDG GTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN
SAGSISANATLTVLETSPFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTER
HFFAAGNQLLIIVDSVDSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDG
WATVGVI IAVVCCVVGTSVLVWVVI IYHTRRRNEDCSITNTDETNPADIPSYLSSQGTLD
RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP
MYLKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISW
PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLD FSANPEPASVASSNSFMGTFGKALRRPHLDA
YSSFGQPSDCQPRAFYLKAHSSPDLD SGSEEDGKERTDFQEENHICTFKQTLNRYRTPNFQS
YDLDT

Signal sequence:

amino acids 1-19

Transmembrane domain:

amino acids 746-765

N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

FIGURE 103

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAATG
AAGGATGCAGGACGCAGCTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGAAC
GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAGAATGAAA
TAAACCAGAGTTAGACCCGCGGGGGTGGTGTGTTCTGACATAAAATAAATAATCTTAAAGCAGCTGTTCCCTCC
CCACCCCCAAAAAAGGATGATTGGAATGAAGAACCGAGGATTACAAAGAAAAAAGTATGTTTCAATTTTTCTC
TATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTTGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT
GGTGTGGTGGTGTTCCTTTCTTTTTGAATTTCCACAAGAGGAGAGGAAATTAATAATACATCTGCAAAGAAA
TTTCAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCACAGTTGGA
TTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTCTTCATCAACCTCCTTTTTTTTAAAT
TTTTATTCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTCTTAACCACCTGGATTTCCATCTGGATGTTGCT
GTGATCAGTCTGAAATACAATGTTTGAATTCAGAAGGACCAACACCAGATAAATTATGAATGTTGAACAAGAT
GACCTTACATCCACAGCAGATAATGATAGGTCTAGGTTTAACAGGGCCCTATTTGACCCCTGCTTGTGGTGCT
GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA
CCAGTTTCAAGGTGATTTGTGTTTCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT
GCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCACTTGGAAATCCT
ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTTGGGGCTTTCAATGGTCTGGCGAAACCTCAACACTCTGGA
ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAACTGAAGGAGCTCTGGTT
GCGAAACAACCCCATTTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCCTTCTTTGCGCCGACTAGACTTAGG
GGAATTGAAAAGACTTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACCTTGAGGTATTTGAACCTTGC
CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAACTAGATGAGCTGGATCTTTCTGGGAATCA
TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAAGTGTGGATGATACAGTCCAGAT
TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACACAATAATCTAAC
ATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACATTTACATCACAACCCTTGGAAGTG
TAACCTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTGTGCCCGGTG
TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT
GATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAGCTGAAATGTGCGGGCCTCCACATC
CCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTGCGGATAGCTGT
GCTCAGTGATGGTACGTTAAATTTCACAAATGTAACCTGTGCAAGATACAGGCATGTACACATGTATGGTGAGTAA
TTCCGTTGGGAATACTACTGCTTCAGCCACCTGAATGTTACTGCAGCAACCACTACTCCTTTCTCTTACTTTTC
AACCCTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCCTCC
AGTGGTTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTTCGACAGAGAAAACCTT
CACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGTGATGAAGACTACCAAATCATCAT
TGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCTGGTCAATTTCTACAAGATGAGGAAGCAGCACCA
TCGGCAAACCATCACGCCCCAACAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC
CATGGAAGCCACCTGCCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT
CAACCACACAACAACAGTTAAACACAATAAATTCATACACAGTTTCAAGTGCATGAACCGTTATTGATCCGAATGAA
CTCTAAAGACAATGTACAAGAGACTCAAATCTAAACATTTACAGAGTTACAAAAACAAACAATCAAAAAA
GACAGTTTATTAAAAATGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAACAA
AAAAGAAAAGAAATTTATTTATTAATAAATTTCTATTGTGATCTAAAGCAGACAAAA

U.S. GOVERNMENT PRINTING OFFICE: 1977

amino acids 1-44

amino acids 523-543

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438,
442-446, 488-492, 606-610

amino acids 183-187

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
391-397, 422-428, 433-439, 531-537

FIGURE 105

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTTCCCTCCTGCTGTTTGGGGGGCA
 TGAAAGGGCTTCGCCGCCGGGAGTAAAAGAAAGGAATTGACCGGGCAGCGCGAGGGAGGAGCGCGCACGCGACCGC
 GAGGGCGGGCGTGCACCCCTCGGCTGGAAGTTTGTGCCGGGCCCCGAGCGCGCGCCGGCTGGGAGCTTCGGGTAGA
 GACCTAGGCCGCTGGACCCGCGATGAGCGCGCCGAGCCTCCGTGCGCGCGCCGCGGGGTGGGGCTGCTGCTGTGC
 GCGGTGCTGGGGCGCGCTGGCCGGTCCGACAGCGCGGTGCGCGGGAACTCGGGCAGCCCTCTGGGGTAGCCGCC
 GAGCGCCCATGCCCACTACCTGCCGCTGCCTCGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGTCTT
 CCCGAGCCACTCCCGTCTGGGTGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTCATCAAGGCAAGTTCC
 ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCA
 GTCTCGGCAAATATTACACTTCTCTCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG
 TTTCACTCCCTTGAACCTTTGGACCTTAGCAGCAACAATATTTTCAGAGCTCCAACTGCATTTCCAGCCCTACAG
 CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACA
 CTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAACTGCCCAACTGCAA
 CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCT
 CTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTTGGGGGCTGAGCAACATGGAAATTTTG
 CAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCTGCAGGAACCTTCAT
 CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA
 ACTTTCATCACTTATCAAGGTTAGATTCAAGCTTCTTGGCCTAAGCTTACTAAATACTGCACATTGGG
 AACACAGAGTCAGCTACATTGCTGATTGTGCTTCCGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAAC
 AATGAAATTTCTGGACTATTGAAGACATGAATGGTGCTTCTCTGGGCTTGACAACTGAGGCGACTGATACTC
 CAAGGAAATCGGATCCGTTCTATTACTAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT
 GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAACTGCAACAATTGCATTTAAATACA
 TCAAGCCTTTTGTGCGATTGCCAGCTAAAAAGGCTCCACAGTGGGTGGCGGAAACAACTTTTCAGAGCTTTGTA
 AATGCCAGTTGTGCCCATCTCAGCTGTCTAAAGGAAGCAATTTTGTCTGTAGCCAGATGGCTTTGTGTGT
 GATGATTTTCCCAAACCCAGATCACGGTTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTC
 ATCTGCTCAGCTGCCAGCAGCAGTGATTCCCAATGACTTTTGCTTGGAAAAAGACAATGAACTACTGCATGAT
 GCTGAAATGGAAATATGACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCCTTCGGCTG
 CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCTCAATCACTTTGGTTTCATCTACTCTGTC
 AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTACCAAGACCCCATGGATCTCACCATCCGAGCTGGGGCC
 ATGGCACGCTTGGAGTGTGCTGTGTGGGGCACCCAGCCCCAGATAGCTTGGCAGAAGGATGGGGGCACAGAC
 TTCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCCGAGATGACGTGTTCTTTATCGTGGATGTGAAGACA
 GAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTTCAGCAAATGCAACTCTGACTGTC
 CTAGAAACACCATCATTTTTCGCGCCACTGTTGGACCGAAGTGAACCAAGGGAGAAACAGCCGCTCTACAGTGC
 ATTTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTTGGTGGTAACCGAGAGGCAC
 TTTTTCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAAATACACATGTGAG
 ATGTCTAACACCCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTATCCCCACTCCAACCTGCGACTCCCCT
 CAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGTGT
 GTGGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGATTGCAGCATTACC
 AACACAGATGAGACCAACTTGGCAGCAGATATTCCTAGTTATTTGTCTCATCTCAGGGAACGTTAGCTGACAGGCAG
 GATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACAGTTTGTACATCTTCAGGTGCTGGATTTTTCTTACCA
 CAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTT
 CTTTGTCCGTTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGAAATGTGTATGGCTCAGATCCTTTTGAACA
 TATCATACAGGTTGCAGTCTTGACCCAAGAACAGTTTAAATGGACCACTATGAGCCAGTTACATAAAGAAAAAG
 GAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT
 GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC
 TCTTTAGATTTTAGTGAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATCTTTCATGGGTACCTTTGGAAAA
 GCTCTCAGGAGACCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTATCCCCACTCCAACCTGCGACTCCCCT
 TTGAAAGCTCATTCTTCCCAGACTTGGACTCTGGGTGAGAGGAAGATGGGAAAGAAAGGACAGATTTTCAGGAA
 GAAATCAGATTTGTACCTTTAAACAGACTTTAGAAAACACAGGACTCCAAATTTTCAGTCTTATGACTTGGAC
 ACATAGACTGAATGAGACCAAAGGAAAGCTTAACTACTACCTCAAGTGAAGTTTATTTAAAGAGAGAGAAT
 CTTATGTTTTTTAAATGGAGTTATGAATTTTAAAGGATAAAAAATGCTTTATTTATACAGATGAACCAAATAC
 AAAAGTTATGAAATTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTATTTTTTAACCTTG
 TTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGATTTTTTATAATGCCAGA
 TTTCTTTTTATGGAAATGAGTTACTAAAGCATTTAAATAATACCTGCCTGTACCATTTTTTAAATAGAAGTT
 ACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

U990401.05

FIGURE 106

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGRGELGQPSGVAAERPCPTTCRCLGDLLDCSR
KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT
LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY
FDNLANTLLVLKLNRRNRI SAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKMQR
NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC
QKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS
WTIEDMNGAFSGLDKLRRLLIQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK
KLQQLHLNTSSLLCDCQLKWLPQWVAENNFSQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDF
PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKNELLHDAEMENY AHLRAQGGE
VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA
RLECAAVGHPPAQIAWQKDG GTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA
GSISANATLTVLETSPFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF
FAAGNQLLIIVDSVDSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDD GWA
TVGVVIIAVVCCVVGTSLVVVIIYHTRRRNEDCSITNTDETNPADIPSYLSSQGT LADRQ
DGYVSSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGPMY
LKG NVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISWPS
HVRKLLNTSYSHNEGPGMKNLCLNKSSLD FSANPEPASVASSNSFMGTFGKALRRPHLDAYS
SFGQPSDCQPRAFYLKAHSSPDLD SGSEEDGKERTDFQEENHICTFKQTL ENYRTPNFQSYDLDT

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 808-828

N-glycosylation site.

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.

amino acids 886-890

Casein kinase II phosphorylation site.

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,
1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.

amino acids 667-675

N-myristoylation site.

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,
513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.

amino acids 58-80, 65-87

U55044 "041101

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 |
|---|---|---|---|---|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-------|
| 0 | 0 | 1 | 4 | 9 | 16 | 25 | 36 | 49 | 64 | 81 | 100 | 121 | 144 | 169 | 196 | 225 | 256 | 289 | 324 | 361 | 400 | 441 | 484 | 529 | 576 | 625 | 676 | 729 | 784 | 841 | 900 | 961 | 1024 | 1089 | 1156 | 1225 | 1296 | 1369 | 1444 | 1521 | 1600 | 1681 | 1764 | 1849 | 1936 | 2025 | 2116 | 2209 | 2304 | 2401 | 2500 | 2601 | 2704 | 2809 | 2916 | 3025 | 3136 | 3249 | 3364 | 3481 | 3600 | 3721 | 3844 | 3969 | 4096 | 4225 | 4356 | 4489 | 4624 | 4761 | 4900 | 5041 | 5184 | 5329 | 5476 | 5625 | 5776 | 5929 | 6084 | 6241 | 6400 | 6561 | 6724 | 6889 | 7056 | 7225 | 7396 | 7569 | 7744 | 7921 | 8100 | 8281 | 8464 | 8649 | 8836 | 9025 | 9216 | 9409 | 9604 | 9801 | 10000 |

CAAAACCTTGCCTCGCGGAGAGCGCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC
 TGGGGGAGCGCGTTTCGGCCTGTGGGGCGCCGCTCGGCGCCGGGGCGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC
 CTGCTCCACGAGGCGCCACTGGTGTGAACCGGGAGAGCCCTGGGTGGTCCCGTCCCCCTATCCCTCCTTTATATA
 GAAACCTTCCACACTGGGAAGGCAGCGCGAGGCAGGAGGGCTCATGGTGAACAAGGAGGCCGGCTGATCTGCAG
 GCGCACAGCATTTCCGAGTTTACAGATTTTACAGATACCAAATGGAAGGCGAGGAGGCAGAACAGCCCTGCCTGGT
 TCCATCAGCCCTGGCGCCGAGGCGCATCTGACTCGGCACCCCTGCGAGGCACCATGGCCCCAGAGCCGGGTGCTGC
 TGCCTCTGCTGCTGCTGCGCCACAGCTGCACCTGGGACCTGTGCTTGGCGTGAGGGCCCGAGGATTTGGCCGAA
 GTGGCGGCCACAGCCTGAGCCCCGAAGAGAACGAATTTGCGGAGGAGGAGCCGGTGTGTTACTGAGCCCTGAGG
 AGCCCCGGGCTGGCCAGCCGCGGTGAGTGCCTCCCGGAGACTGTGCTGTTCAGGAGGGCGTGTGGACTGTG
 GCGGTATTGACCTGCGTGAGTTCCCGGGGGACCTGCCTGAGCACACCAACCACCTATCTCTGCAGAACAAACAGC
 TGGAAAAGATCTACCTGAGGAGCTCTCCCGGCTGCACCGGCTGGAGACACTGAACCTGCAAAACAACCGCCTGA
 CTTCCCGAGGGCTCCAGAGAAGGCGTTGTGACATCTGACCAACCTCAATTACCTGTACTTGGCCAATAACAAGC
 TGACCTTGGCACCCCGCTTCTGCCAAACCGCCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCAACAGATCT
 ATGGGCTCACCTTTGGCCGAGAAGCCAAACTTGAGGTCTGTGTACTGCAACAACAAGCTGGCAGACGCCGGGC
 TGCCGGACAACATGTTCAACGGCTCCAGCAACGTGAGGTCTCATCTGTCCAGCAACTTCTGCGCCAGTGC
 CCAAGCACCTGCCGCTGCCCCTGTACAAGCTGCACCTCAAGAACAACAAGCTGGAGAAGATCCCCCGGGGGCCT
 TCAGCGAGCTGAGCAGCCTGCGCGAGCTATACCTGCAGAACAACTACCTGACTGACGAGGGCCTGGACAACGAGA
 CCTTCTGGAAGCTCTCCAGCCTGGAGTACCTGGATCTGTCCAGCAACAACCTGTCTCGGTCAGCTGGGCTGC
 CGTCAGCCTGGTGTGCTGTGCATTTGGAGAAGAAGCCACTCCGGAGCGTGGACGCGAATGTGCTGACCCCATCC
 GCAGCCTGGAGTACCTGCTGTGCTGCACAGCAACCAGCTGCGGGAGCAGGGCATCCACCACCTGGCCTTCCAGGGCC
 TCAAGCGGTTGCACACGGTGCACCTGTACAACAACGCGCTGGAGCGCGTGCCAGTGGCCTGCCTCGCCGCGTGC
 GCACCCTCATGATCTTGCACAACCAGATCACAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCTTGGAGG
 AGCTCAACCTCAGCTACAACCGCATCACAGCCCAAGGTGCACCGCGACGCCCTCCGCAAGCTGCGCCTGCTGC
 GCTCGCTGGACCTGTGCGGCAACCGGCTGCACAGCTGCCACCTGGGCTGCCCTCGAAATGTCCATGTGTGGAAG
 TCAAGCGCAATGAGCTGGCTGCCCTTGGCAGAGGGGCGCTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACTCA
 CCAAGCAACGACTGCGCAGCCGAGCCCTGGGCCCCCGTGCCTGGTGGACCTCGCCCATCTGCAGCTGCTGGACA
 TCGCCGGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACACA
 AGATTAGTGCAGTGCCTGCCCAATGCCCTTGCACCTCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAAACAAGC
 TGGCTGTGGGCTCCGTGGTGGACAGTGCCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGGCAACT
 TAGAGTTTGGTGACATTTCCAAGGACCGTGGCCGCTTGGGGAAGGAAAAGGAGGAGGAGGAGGAGGAGGAGG
 AGGAAGAGGAAAACAAGATAGTGACAAGGTGATGCAGATGTGACCTAGGATGATGAGCCCGGACTCTTTCTGC
 AGACACGCTGTGTGCTGTGAGCCCCCACTCTGCGGTGCTCACACAGACACACCAAGCTGCACACATGAGGCA
 TCCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCCTTCCACGGCGTGTCCACGGCCAGACACATGC
 ACACACATCACACCCTCAAACACCCAGCTCAGCCACACACAACCTACCTCCAAACCACCACAGTCTCTGTACAC
 CCCACTACCGCTGCCACGCCCTCTGAATCATGCAGGGAAGGGTCTGCCCTTGCCTTGGCACACACAGGCACCCA
 TTCCCTCCCCCTGCTGACATGTGTATGCGTATGCATACACACACACACACACATGACACAGTCAATGTGCGAA
 CAGCCCTCAAAGCCTATGCCACAGACAGTCTTTGCCCGAGCCAGAATCAGCCATAGCAGCTCGCCGTCTGCCCT
 GTCCATCTGTCCGTCCGTTCCCTGGAGAGAAGACACAAGGGTATCCATGTCTGTGTGGCCAGGTGCCTGCCACCTCT
 GGAACCTCAAAAAGCTGGCTTTTATTCTTTCCCATCCTATGGGGACAGGAGCCTTCAGGACTGCTGGCCTGGCC
 TGGCCACCCCTGCTCCTCCAGGTGCTGGGCAGTCACTCTGCTAAGAGTCCCTCCCTGCCACGCCCTGGCAGGACA
 CAGGCACCTTTTCCAATGGGCAAGCCAGTGGAGGCAGGATGGGAGAGCCCCCTGGGTGCTGTGGGGCCTTGGGG
 CAGGAGTGAAGCAGAGGTGATGGGGCTGGGCTGAGCCAGGGAGGGAAGGACCCAGCTGCACCTAGGAGACACCTTT
 GTTCTTCAAGCCTGTGGGGGAAGTTCGGGTGCCCTTATTTTATTTCTTTTCTAAGGAAAAAATGATAAAAAAT
 CTCAAAGCTGATTTTCTTGTATAGAAAAATTAATATAAAGCATTTATCCCTATCCCTGCAAAAAAATAA

FIGURE 108

MEGEEAEQPAWFHQPWPRPGASDSAPPAGTMAQSRVLLLLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSPREENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVDCGGIDLREFFPGDLP
EHTNHLSLQNNQLEKIYPEELSRLETLNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVLDANVLTPIRSLEYLLHSN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMLHNQITGIGREDFATTYF
LEELNLSYNRITSPQVHRDAFRKLRLRLSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAAAL
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKD
RGRLGKEKEEEEEEEEEEEEEETR

Signal sequence:

amino acids 1-48

N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,
477-483, 498-502, 539-545, 548-554

Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,
535-557

FIGURE 109

GGGAGGGGGCTCCGGGCGCCGCGCAGCAGACCTGCTCCGGCCGCGCGCCTCGCCGCTGTCCTCCGGGAGCGGCAG
CAGTAGCCCCGGGCGGCGAGGGCTGGGGGTTCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCCCACCAACC
CAACCTGTTCTCGCGCGCCACTGCGCTGCGCCCCAGGACCCGCTGCCCAACATGGATTCTCTCTGGCGCTGGT
GCTGGTATCTCTCGCTTACCTGTCAGGCGGCGCCGAGTTTCGACGGGAGGTGGCCAGGCAAATAGTGTATCGAT
TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTAGCC
TGTGTGCCAACACCGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTATCCTGGTTATGCTGG
AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA
CGGCAGCTACAAGTGTACTGTCTCAACGGATATATGCTCATGCCGATGGTTCTGCTCAAGTGCCCTGACCTG
CTCCATGGCAAACCTGTAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCCTGGCCT
GCACCTGGCTCTGTATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT
TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTATATAAGGCTTCGATCTCATGTATATTGGAGG
CAAATATCAATGTTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA
CGTACGTGGGTCTTACAAGTGCAAATGTAAAGAAGGATACCAGGGTGATGGACTGACTTGTGTGTATATCCCAAA
AGTTATGATTGAACCTTCAGGTCCAATTTCATGTACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAA
TAATAATTGGATTCTGTATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAA
CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCTACTCCACCACCACCACCACC
CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCAGAAAGGCCAACACCACCGGACTGACAACTATAGC
ACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCCCTCAGAAACCCAGAGG
AGATGTGTTTCTGTTTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAA
TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCCGCAGCCAAAGCCCCAGG
GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCCTCATGCATTAGGGGACCTGTGCCTGTATTAGGCA
CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAACACGGTGCCACGGAGCAGCCTGTG
GGGAAGAAATGGTGGCCATGGCTGGAGGCAAAACACAGATCACCTTGCGAGGGGCTGACATCAAGAGCGAATCACA
AAGATGATTAAAGGGTTGGAAAAAAGATCTATGATGGAATAAAGGAACTGGGATTATTGAGCCTGGAGAAG
AGAAGACTGAGGGGCAAACCATTTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG
TTCTCCATATGCACTAAGAATAGAACAAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGCAGG
GGCCATTGTTAGAATACTTCATAAAAAAAGAGTGTGAAATCTCAGTATCTCTCTCTTTCTAAAAAATTAGA
TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCCAAA
AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG
CATTTCCCTCTTGACCTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCCACCAAATGCTGAGCTCACTGAAATA
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAAGGAACTATTTATTCCAAATGAGAGTATGATGGAC
AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC
TTTCATTTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTTTAAGAGATCCTTCAAGGAACACAGTTTCAGAGAG
ATTTTCATCGGGTGCACTCTCTCTGCTTCGTGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC
ACACCGGCAGACCTTTCTTTCACCTCATCAGTATGATTAGTTTCTCTTATCAATTGGACTCTCCCAGGTTCCAC
AGAACAGTAATATTTTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCAATTAACCTGGTAAAGGCAGGGCTGG
AGGGGGAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTAAATGGTTTCATT
TCCTTTATGGTCATATAACTGCACAGCTGAAGATGAAAGGGGAAAAATAAATGAAAATTTTACTTTTCGATGCCAA
TGATACATTGCACTAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTATTATTATTTAATGTTTT
CTAAAAATAAAAAATGTTAGTGGTTTTTCCAAATGGCCTAATAAAAAACAATTATTTGTAAATAAAAAACTGTTAGTAAT

FIGURE 110

MDFLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQP
RCKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD
GSCSSALTCSMANCQYGC DVVKGQIRCQCPSGLHLAPDGRTCV DVDECATGRASCPRFRQC
VNTFGSYICKCHKGF DL MYIGGKYQCHDIDEC SLGQYQCSSFARCYNVRGSYKCKCKEGYQG
DGLTCVYIPKVMIEPSGPIHVPKGN GTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRP
TSKPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDN
RVQTD PQKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGG
KAARLVLP LGR L MHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQI
TLRGADIKSESQR

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

N-myristoylation site.

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,
421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

FIGURE 111

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCCTTTAGATTGTGA
AATGTGGCTCAAGGTCTTCACAACCTTTCTTTCTTTGCAACAGGTGCTTGCTCGGGGCTGA
AGGTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGTC
CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA
CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCCTGACTTGGAATACC
AACACAAGTTTACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT
GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACCTATCTGCCAGTCA
GAAGATACAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTCATCCTCCCT
CTGGGGCTGTGGAGTATGTGGGGAACATGACCTTGACATGCCATGTGGAAGGGGGCACTCGG
CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC
TCCCCAAAACAATAACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT
GCCTGGTGAGGAACCTGTGAGTGAATGGAAAGTGATATCATTATGCCCATCATATATTAT
GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT
TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACT
CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTT
GCATCTGAGAAAGTAGCCGAGAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC
CGGCAGGCAAGATGAACTCATTTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG
CACAGAAAGGAAAATCATTGTACCTTTAGCAAGTATAACTGGAATATCACTATTTTTTGATT
ATATCCATGTGTCTTCTCTTCTTATGGAAAAATATCAACCCTACAAAGTTATAAAACAGAA
ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG
CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTCAGG
ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT
GTATGAAGTTATTCAGCACATCCCTGCCAGCAGCAAGACCATCCAGAGTGAACCTTTCATGG
GCTAAACAGTACATTCGAGTGAAATTCCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT
ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA
TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT
GTGCAACAGAAAAACATGTTGGGGAAATATTCTCAGTGGAGAGTCGTTCTCATGCTGACGG
GGAGAACGAAAGTGACAGGGGTTTCCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA
ATTAGTTCTACTCTACACTTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA
TGTGGAACTTTACATTGTTTCGATTTTTCAGCAGACTTTGTTTTATTAAATTTTATTAGTG
TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA
CAAAGTAATAAGGATGGTTGTCACAAAAACAAAACCTATGCCTTCTCTTTTTTTCAATCACC
AGTAGTATTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTTA
TTTTTTTCAAGGAAAGATGGATTCAAATAAATTATTCTGTTTTTGCTTTTAAAAA

FIGURE 112

MWLKVFTTFLSFATGACSGLKVTVPSTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH
TMPKYLLGSVNKSVVPDLEYQHKFTMMPNASLLINPLQFPDEGNYIVKVNIQGNGTLSASQ
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNRGPVHTSSTYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT
GRQDETHFTVIITSVGLEKLAQKGKSLSPLASITGISLFLIISMCLLFLWKKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRVSPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208,
276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237,
239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 113

GCAAGCGGCGAAATGGCGCCCTCCGGGAGTCTTGCAAGTCCCTGGCAGTCCTGGTGCTGTT
GCTTTGGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTTCGCGTCATCACGGACGAGA
ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCT
TGTCAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAA
TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC
TTCCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG
AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCATG
GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGA
TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT
GTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC
AGATTGCCTTTGTCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCTTCAAAAAAAT
TATTATCAGAATCTGCACAACCTTTGAAAAAAGTGAGGAGGAACAAGAGCGGATGAAGAA
GATGTTTCAGAAGAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTTTCACAGAATGC
CATAAGACAACGCTCTCTGGGTCCATCATTGGCCACAGATAAATCCTAGTAAATTTTATAG
TTATCTTAATATTATGATTTTGATAAAAAACAGAAGATTGATCATTTTGTGTTGGTTGAAGTG
AACTGTGACTTTTTTGAATATTGCAGGGTTCAGTCTAGATTGTCAATTAAATTGAAGAGTCTA
CATTCAGAACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG
GTTTAAATAGTTCTCTAATTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGT
TTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAAATTTACATTTCCCAAGTATTGCATTAT
TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTG
TTTCACTGTGTGAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATTTGTCTCAAG
AAATGTGTATTTCACTGACAATTTCTGTTGCTTTTTTAGAGGTATATTCCAAAATTTCTTGT
ATTTTTTAGGTTATGCAACTAATAAAAACTACCTTACATTAATTAATTACAGTTTTCTACACA
TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTTAAGTTCATGGTATTCTCTTGATTC
CAACAAAGTTTGATTTTCTCTTGATTTTTTCTTACTTACTATGGGTACATTTTTTATTTTT
CAAATTGGATGATAATTTCTTGAAACATTTTTTATGTTTTAGTAAACAGTATTTTTTTGTT
GTTTCAAACCTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAATTT
TTGGCCACTTTTTTTCAGATTTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTT
TTTCTTTTTTGGATGTGAAGGTGAACATTCCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTA
TTTTACATTTTTGAAAATTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAG
CATCTTCTTGATATATGTCTTAAATGTATTTTTGTCTCATATACAGAAAGTTCTTAATTGAT
TTTACAGTCTGTAATGCTTGATGTTTTAAATAATAACATTTTTATATTTTTTAAAGACAA
ACTTCATATTATCCTGTGTTCTTTCCTGACTGGTAATATTGTGTGGGATTTACAGGTAATA
GTCAGTAGGATGGAACATTTTAGTGTATTTTTACTCCTTAAAGAGCTAGAATACATAGTTTT
CACCTTAAAAGAAGGGGGAAAATCATAAATAACAATGAATCAACTGACCATTACGTAGTAGAC
AATTTCTGTAATGTCCCCTTCTTTCTAGGCTCTGTTGCTGTGTGAATCCATTAGATTTACAG
TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAATATTGTACCATT
AAAGAGTTTGGATGTGTAACCTGTGATGCCTTAGAAAAATATCCTAAGCACAAAATAAACCT
TTCTAACCACTTCATTAAAGCTGAAAAAATAAAAAAAAAA

FIGURE 114

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL
QPEWESFAEWGEDLEVNIKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI
NFISDKEWKSIIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL
ATLFSGLLLGLCMIFVADCLCPSKRRRPQYPYPYPSKLLSESAQPLKKVEEQEAEDEEDVSE
EEAESKEGTNKDFPQNAIQRSLGPSLATDKS

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

FIGURE 115

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTAACTAATTCAACAAACGGGACCCTT
CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCAGTGGGACAGGCGGATTGGAAGAGCGGG
AAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAAACTCTGGGTGTCTGC
ATTGCTGATGGCCTGGTTTGGTGTCTGAGCTGTGTGCAGGCCGAATTCTTCACCTCTATTG
GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC
CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC
TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCTGTGAATGCCTACAAACTGG
TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTGCAGGACTCAGCTGCA
GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGG
AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA
GAGGGGAACCTTCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG
GGCCGCTCGGCCTACAATGAAGGGGACTATTATCATAACGGTGTGTGGATGGAGCAGGTGCT
AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCA
GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTC
TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGA
GGAAGAGAGAGAAAAACGTTAAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAGGCA
TCTATGAGAGGCCTGTGGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCTGGG
GAGGGTGTCAAACCTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA
CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACA
TCGTACAGGTACTACGATGTCATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA
CCTAAACTTGCACGAGCCACCGTTCTGTGATCCCAAGACAGGAGTCCTCACTGTGCCGACTA
CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATC
GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT
TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG
CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG
GTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG
TTCTGGTACAACCTCTTGCGGAGCGGGGAAGGTGACTACCGAACAAGACATGCTGCCTGCCC
TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT
TGAGACCTTGTGGATCAACAGAAGTTGACTGACATCCTTTTCTGTCTTCCCCTTCCTGGTC
CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCTTTGTATGTTTCTATCAGGCT
GATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT
GTGACTGAAGTCCAGCCCTTCCATTACAGCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA
AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTTGTACCTCAGGTGTT
TTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTTGTTTTAT
GGCATTCTATCTATTGTGGCTTTACCAAAAAATAAAATGTCCCTACCAGAAAAA

FIGURE 116

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA
NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALDVLQDSAAGFIANLSVQRQFFP
TDEDEIGAALKMRLQDQTYRLDPGTISRGLPGTKYQAMLSVDDCFGMGRSAYNEGDDYYHTV
LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLSLDPSHERAGGNLR
YFEQLLEEEREKTLTNQTEAELATPEGIYERPVDYLPERDVYESLCRGEGVKLTTPRRQKRLF
CRYHHGNRAPQLLIAPFKEEDEWDSPIVRYDDVMSDEEIERIKEIAKPKLARATVRDPKTG
VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS
RRPFDSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGECDYR
TRHAACPVLVGCKWVSNKWFHERGQEFRLPCGSTVD

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,
346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

FIGURE 117

GCAGTATTGAGTTTTACTTCCTCCTCTTTTTAGTGGAAGACAGACCATAATCCCAGTGTGAGTGAAATTGATTGT
TTCATTTATTACCGTTTTGGCTGGGGGTAGTTCCGACACCTTCACAGTTGAAGAGCAGGCAGAAGGAGTTGTGA
AGACAGGACAATCTTCTTGGGGATGCTGGTCCTGGAAGCCAGCGGGCCTTGCTCTGTCTTTGGCCTCATTGACCC
CAGGTTCTCTGGTTAAAACTGAAAGCCTACTACTGGCCTGGTGCCCATCAATCCATTGATCCTTGAGGCTGTGCC
CCTGGGGCACCCACCTGGCAGGGCCTACCACCATGCGACTGAGCTCCCTGTTGGCTCTGCTGCGGCCAGCGCTTC
CCCTCATCTTAGGGCTGTCTCTGGGGTGACGCTGAGCCTCCTGCGGGTTTCTTGATCCAGGGGGAGGGAGAAG
ATCCCTGTGTGAGGCTGTAGGGGAGCGAGGAGGGCCACAGAATCCAGATTTCGAGAGCTCGGCTAGACCAAAGTG
ATGAAGACTTCAAACCCCGGATTGTCCCCTACTACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTC
GGTACATCCAGACAGAGCTGGGCTCCCGTGAGCGGTTGTGGTGGCTGTCTGACCTCCCAGCTACACTGTCCA
CTTTGGCCGTGGCTGTGAACCGTACGGTGGCCCATCACTTCCCTCGGTTACTCTACTTCACTGGGCAGCGGGGGG
CCCGGGCTCCAGCAGGGATGCAGGTGGTGTCTCATGGGGATGAGCGGCCCCGCTGGCTCATGTGAGAGACCTGC
GCCACCTTCACACACACTTTGGGGCCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGCAGGCCC
CCCGCCTGGCAGCCCTTGCTGGCCACCTCAGCATCAACCAAGACCTGTACTTAGGCCGGGCAGAGGAGTTCATTG
GCGCAGGCGAGCAGGCCCGGTACTGTCTATGGGGCTTTGGCTACCTGTTGTACGAGAGTCTCCTGCTTCGTCTGC
GGCCACATCTGGATGGCTGCCGAGGAGACATTCTCAGTGCCCGTCTGACGAGTGGCTTGGACGCTGCCTCATTG
ACTCTCTGGGCGTCGGCTGTGTCTCACAGCACCAGGGGCGAGCAGTATCGCTCATTTGAAGTGGCCAAAAATAGGG
ACCCTGAGAAGGAAGGGAGCTCGGCTTTCTGAGTGCCTTCGCCGTGCACCCCTGTCTCCGAAGGTACCCTCATGT
ACCGGCTCCACAAACGCTTCAGCGCTCTGGAGTTGGAGCGGGCTTACAGTGAAATAGAACAACTGCAGGCTCAGA
TCCGGAACCTGACCGTGCTGACCCCCGAAGGGGAGGCAGGGCTGAGCTGGCCCCGTTGGGCTCCCTGCTCCTTTCA
CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACACCTTCTCCTGTGCAGATGGGG
CTCCCAAGTGCCCACTACAGGGGGCTAGCAGGGCGGACGTGGGTGATGCGTTGGAGACTGCCCTGGAGCAGCTCA
ATCGGCGCTATCAGCCCCGCTGCGCTTCCAGAAGCAGCGACTGCTCAACGGCTATCGGCGCTTCGACCCAGCAC
GGGGCATGGAGTACACCTGGACCTGCTGTTGGAATGTGTGACACAGCGTGGGCACCGGCGGGCCCTGGCTCGCA
GGGTGAGCCTGCTGCGGCCACTGAGCCGGGTGGAATCCTACCTATGCCCTATGTCACTGAGGCCACCCGAGTGC
AGCTGGTGCTGCCACTCCTGGTGGCTGAAGCTGCTGCAGCCCCGGCTTTCTTCGAGGCGTTTGCAGCCAATGTCC
TGGAGCCACGAGAACATGCATTGCTCACCCCTGTTGCTGGTCTACGGGCCACGAGAAGGTGGCCGTGGAGCTCCAG
ACCCATTTCTTGGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGGTACCCTGGGACGAGGCTGGCCTGGCTCG
CTGTGCGAGCAGAGGCCCTTCCCAGGTGCGACTCATGGACGTGGTCTCGAAGAAGCACCCCTGTGGACACTCTCT
TCTTCCTTACCACCGTGTGGACAAGGCCTGGGCCCGAAGTCTCAACCGCTGTGCGCATGAATGCCATCTCTGGCT
GGCAGGCCTTCTTTCCAGTCCATTTCAGGAGTTCAATCCTGCCCTGTACCCACAGAGATCACCCCCAGGGCCCC
CGGGGGCTGGCCCTGACCCCCCTCCCTCCTGGTGTGACCCCTCCCGGGGGGCTCCTATAGGGGGGAGATTTG
ACCGGCAGGCTTCTGCGGAGGGCTGCTTCTACAACGCTGACTACCTGGCGGGCCGAGCCGGGCTGGCAGGTGAAC
TGGCAGGCCAGGAAGAGGAGGAAGCCCTGGAGGGGCTGGAGGTGATGGATGTTTCTCCGGTTCTCAGGGCTCC
ACCTCTTTCCGGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGCAGCCACGGCTCAGTGAAG
AACTCTACCACCGCTGCCGCCTCAGCAACCTGGAGGGGCTAGGGGGCCGTGCCAGCTGGCTATGGCTCTCTTTG
AGCAGGAGCAGGCCAATAGCACTTAGCCCCCTGGGGGCCCTAACCTCATTACCTTTCTTTGTCTGCCTCAGCC
CCAGGAAGGGCAAGGCAAGATGGTGGACAGATAGAGAATTGTTGCTGTATTTTTTAAATATGAAAATGTTATTAA
ACATGTCTTCTGCC

U
S
S
R
I
N
S
T
I
T
U
T
E

FIGURE 118

MRLSSLLALLRPALPLIILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPDSRARLD
QSEDFKPRIVPYRDPNKPYPKKVLRTRYIQTELGSRRERLLVAVLTSRATLSTLAVAVNRTV
AHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYDWFFIMQDDTY
VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRSLLLRLRPHLDGCRG
DILSÄRPDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE
GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPEGEAGLSWPVGLPAPFTPHSRFEV
LGWDYFTEQHTFSCADGAPKCPLOGASRADVGDALETALEQLNRRYQPRLRFQKQRLNNGYR
RFDPARMEYTLDLLLEC VTQRGHRRALARRVSLLRPLSRVEILPMPYVTEATRVQLVLP LL
VAEAAAAPAFLEAFAANVLEPREHALITLLL VYGPREGGRGAPDPFLGVKAAAELERRY PG
TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLT TVWTRPGPEVLNRCRMNAISGWQAFPP
VHFQEFNPALSPQRSPPGPPGAGDPSPSGADPSRGAPIGGRFDRQASAEGCFYNADYLAA
RARLAGELAGQEEEEALEGLEVMDFLRFSGHLHFRAVEPGLVQKFSLRDCSPRLSEELYHR
CRLSNLEGLGGRAQLAMALFEQE QANST

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 489-507

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

FIGURE 119

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCC
GTTCTAGACGCGGGAAAAATGCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG
AAGCATTTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA
TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTCA
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAACC
CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG
AGTTCTTCAGTTCTGAAAATGTTAAAGTGTGTTGAGTCAATTAATATGGACACAAATGACATG
TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACCTG
GTTCTTCCTTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTAA
AAAAGGATCCATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT
GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT
CAATATCCCAGAAAAGTGTCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC
AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA
AAAGATGTATTTAATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC
CAACCAGGTAGTAGAAGGCTGTTGTTTCAGATATGGCTGTTACTTTTAATGGACTGACTCCAA
ATCAGATGCATGTGATGATGTATGGGGTATACCGCCTTAGGGCATTTGGGCATATTTTCAAT
GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG
TGAATATGATCTTTGTATAGGACGTGTGTTGTCATTATTTGTAGTAGTAACATACATATCCAA
TACAGCTGTATGTTTCTTTTTCTTTTCTAATTTGGTGGCACTGGTATAACCAACACATTAAAG
TCAGTAGTACATTTTTTAAATGAGGGTGGTTTTTTCTTTAAACACATGAACATTGTAAATG
TGTTGGAAAGAAGTGTTTTAAGAATAATAATTTGCAAATAAACTATTAATAAATATTATAT
GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTT
AAAAAATTTTAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT
TGTGATTAAAGTAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT
AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAATGTACAACCTAAGCAACTAAAGAAA
ATTAAAGTGAAAGTTGAAAAAT

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCC

FIGURE 120

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISED
LSKSFVYCIILVKPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRK
AYKYAFDKYRDQYNWFFLARPTTFAI IENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGM
IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDV
FNT KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL
PPNGSDND

Signal sequence:

amino acids 1-33

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

U
S
N
A
T
O
N
A
L
B
I
O
T
E
C
H
N
I
C
J
O
U
R
N
A
L
O
F
M
O
L
E
C
U
L
A
R
B
I
O
C
H
E
M
I
S
T
R
Y

FIGURE 121

CCCACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAG
AAAAAGAGAGAGAGAGAGAAACAAAAAACCAAAGAGAGAGAGAAAAAATGAATTCATCTAAATCAT
CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCT
GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTTCGCAT
CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTTACAGAGCTCTCCTGCT
ACAATTATGGATCAGGTTCAAGTCAAGTGTGTTGTCCATTGAACTGGGAATATTTTCAATCC
AGCTGCTACTTCTTTTCTACTGACACCATTTCTGCGGCTTAAGTTTAAAGAACTGCTCAGC
CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCTTACAAGA
AACCTAAAATGAGAGAGTTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA
TGGGTGGACGGCACACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA
CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA
ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG
AACAAAGGAAAATCTCTTTAAAGAACAGAAGGCACAACCTCAAATGTGTAAAGAAGGAAGAGCA
AGAACATGGCCACACCCACCGCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC
ATAAGTATTTGTTACTCTGATACAAATAAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAAAAA
AAA
AAAAA

U.S. G. P. O.

FIGURE 122

MNSSKSSETQCTERGCFSQMFLLTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN
FTELSYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSCLKNCSAMGAHLVVINSQEEQ
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCAIMRDSS
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

Signal sequence:

amino acids 1-42

N-glycosylation site.

amino acids 2-6, 62-66, 107-111

Casein kinase II phosphorylation site.

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

N-myristoylation site.

amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 27-38

U.S. GOVERNMENT PRINTING OFFICE: 1975

FIGURE 123

GGGACTACAAGCCGCGCCGCGCTGCCGCTGGCCCCCTCAGCAACCCCTCGACATGGCGCTGAGGCGGCCACCGCGAC
TCCGGCTCTGCGCTCGGCTGCCCTGACTTCTTCTGCTGCTGCTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC
TCAAATCCAGCAATCGAACCCAGTGTTACAGGAATTTGAAAGTGTGGAAGTGTCTTGCATCATTACGGATTTCGC
AGACAAGTGACCCCAAGGATCGAGTGGAAAGAAAATTCAGATGAACAAACCATATGTGTTTTTTTGACAACAAAA
TTCAGGGAGACTTGGCGGGTCTGTCAGAAATCTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG
ACTCAGCCCTTTATCGCTGTGAGGTCGTTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA
CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC
ACTGCCAGGAGAGTGAGGGCCACCCCGGCTCACTACAGCTGGTATCGCAATGATGTACCCTGCCCACGGATT
CCAGAGCCAATCCCAGATTTTCGCAATCTTCTTTCCACTTAAACTCTGAAACAGGCACCTTTGGTGTTCCTGCTG
TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG
AGATGGAAGTCTATGACCTGAACATTTGGCGGAATTATTGGGGGGGTTCTGGTTGTCTTGTGCTACTGGCCCTGA
TCACGTTTGGGCATCTGCTGATGCATACAGACGTGGCTACTTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA
ACCCAGGGAACAGATGAGGTTAACTTACATCCGCACTGACGAGGAGGGCGACTTCAGACACAAGTCTATCGTTTG
TGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA
GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTGAGAAGCTTTTCGTTTTGGCCAAAGTTGACCA
CTACTCTTCTTACTCTAACAAGCCACATGAATAGAGAATTTTCTCAAGATGGACCCGGTAAATATAACCACAA
GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCTAATCTGTTTTCTGGCCTGATTCCCGCATGAGTATTAGG
GTGATCTTAAAGAGTTTGTCTACGTAAACGCCCGTGTCTGGGCCCTGTGAAGCCAGCATGTTACCCTGGTCTGTT
CAGCAGCCACGACAGCACCATGTGAGATGGCGAGGTGGCTGGACAGCACCAGCAGCGCATCCCGCGGGGAACCCA
GAAAAGGCTTCTTACACAGCAGCCTTACTTTCATCGGCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC
TGATCGGTGTGTCAGTGTCCATTGTGGAGAAGCTTTTGGATCAGCATTTTGTAAAAACAACCAAAATCAGGAAG
GTAAATTGGTTGCTGGAAGAGGGATCTTGCCTGAGGAACCTGCTTGTCCAACAGGGTGTGAGGATTTAAGGAAA
ACCTTCGTCTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTTGTATTATTTATAAAATTT
TACATCTAAATTTTGTCTAAGGATGTATTTGATTATTGAAAAGAAAATTTCTATTAACTGTAAATATATTGT
CATACAATGTTAAATAACCTATTTTTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTAAAT
TGGAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTCTTTTCT
CACACAAGTTTTAGCCTTTTTTACAAGGGAACCTACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT
TAAAAATTCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAACT
GCCTCTTCTGAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGAAGCCCTCAGATGTACATACACAGATG
CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCGCTCTAGCTCACTGTTGCCTCGCTGTCTGCCAGGAGGCCCT
GCCATCCTTGGGGCTGGCAGTGGCTGTGTCCCAGTGAGCTTTACTCACGTGGCCCTTGCTTTCATCCAGCACAGC
TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT
TTTTGGTTATGGATGGCTCACAAAATAGGGCCCCCAATGCTATTTTTTTTTTTAAGTTGTTTAATTATTTGTT
AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTAAAAAGAAAATGGAT
CCCACTGTTTCTCTTTGCCACAGAGAAAGCACCCAGACGCCACAGGCTCTGTGCGATTTCAAAACAAACCATGAT
GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTGAGTGGAGCAGCCAGGTGAAAGGCCTGGCGGGGAGGAAAG
TGAAACGCCTGAATCAAAGCAGTTTTCTAATTTTGACTTTAAATTTTTTATCCGCGGAGACACTGCTCCCAT
TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT
GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCCTGGA
GAATGGCTCTCACTACTCACCTTGTCTTTTCAAGCTTCCAGTGTCTTGGGTTTTTTATACTTTGACAGCTTTTTTTT
AATTGCATACATGAGACTGTGTTGACTTTTTTTAGTTATGTGAAACACTTTGCGCGAGGCCCGCTGGCAGAGGCA
GGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGCTGCATGGCATCCTGGATGCTTAGCATGCAAGTTC
CCTCCATCATTGCCACCTTGGTAGAGAGGGATGGCTCCCCACCCTCAGCGTTGGGGATTACGCTCCAGCCTCCT
TCTTGGTTGTCTAGTGTAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA
TGGGAACCAGGTCTGAAAAAGTAGAGAGAAGTGAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA
CGGAAAAGGAATCTCGTGTATTTTAAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT
GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTTGTACTAACACACCGTAATTTGGCATTGTTTAAAC
CTCAATTTATAAAAGCTTCAAAAAAACCA

TTTCTTAAAGGCTCTGC

Figure 1 displays 12 histograms showing the distribution of the number of non-zero elements in the vector x_k for $k = 0, 1, \dots, 11$. The x-axis represents the number of non-zero elements, ranging from 0 to 100. The y-axis represents the frequency, ranging from 0 to 10. The distributions are roughly bell-shaped and centered around 50, with the peak frequency increasing from $k=0$ to $k=11$.

```
><subunit 1 of 1, 310 aa, 1 stop
```

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTS
PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLKIWNVTRRDSALYRCEVVARNDRK
EIDEIVIELTVQVKPVPVPCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPLPTDSRA
NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVDLNIIGGIIG
VLVVLAVLALITLGICCAYRRGYFINNKODGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

Signal peptide:

Transmembrane domain:

N-glycosylation sites.

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Casein kinase II phosphorylation site.

Tyrosine kinase phosphorylation site.

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267